**JOINT INVENTORS**

"EXPRESS MAIL" mailing label No.
EL564460153US.
Date of Deposit: April 16, 2001
I hereby certify that this paper (or fee) is being deposited with the United States Postal Service "EXPRESS MAIL POST OFFICE TO ADDRESSEE" service under 37 CFR §1.10 on the date indicated above and is addressed to: Commissioner for Patents, Washington, D.C. 20231

Lynn L. Janulis

Lynn Janulis

# APPLICATION FOR UNITED STATES LETTERS PATENT

# SPECIFICATION

**TO ALL WHOM IT MAY CONCERN:**

**Be it known that we, Dennis G. Ballinger a citizen of United States, residing at 38 Bishop Lane, in the City of Menlo Park, 94025 and State of California, Deborah Loeb a citizen of United States, residing at 5932 Valley Meadow Court, in the City of San Jose, 95135 and State of California, Julie R. Montgomery a citizen of United States, residing at 23411 Glenwood Drive, in the City of Los Gatos, 95033 and State of California, Y. Tom Tang a citizen of United States, residing at 4230 Ranwick Court, in the City of San Jose, 95118 and State of California, Ping Zhou a citizen of People's Republic of China, residing at 1461 Japaur Lane, in the City of San Jose, 95132 and State of California, Ryle Goodrich a citizen of United States, residing at 4896 Sandy Lane, in the City of San Jose, 95124 and**

State of California, Chenghua Liu a citizen of United States, residing at 1125 Ranchero Way, Apt. 14, in the City of San Jose, 95117 and State of California, Vinod Asundi a citizen of United States, residing at 709 Foster City Blvd., in the City of Foster City, 94404 and State of California, Qing A. Zhao a citizen of People's Republic of China, residing at 1028 S. de Anza Blvd., Apt. B-210, in the City of San Jose, 95129 and State of California, Tom Wehrman a citizen of United States, residing at 300 Pasteur Drive, Edwards R314, Stanford University Medical Center, in the City of Stanford, 94305 and State of California, Radoje T. Drmanac a citizen of Yugoslavia, residing at 850 East Greenwich Place, in the City of Palo Alto, 94303 and State of California, Feiyan Ren a citizen of People's Republic of China, residing at 20685 Garden Manor Court, in the City of Cupertino, 95014 and State of California, Xiaohong B. Qian a citizen of People's Republic of China, residing at 3662 Tumble Way, in the City of San Jose, 95132 and State of California and Dunrui Wang a citizen of People's Republic of China, residing at 932 La Palma Place, in the City of Milpitas, 95035 and State of California have invented a new and useful MATERIALS AND METHODS RELATING TO LIPID METABOLISM, of which the following is a specification.

# MATERIALS AND METHODS RELATING TO LIPID METABOLISM

This application claims priority of U.S. provisional application no. 60/197,137 filed April 14, 2000, and is a continuation-in-part of US applications serial nos. 09/714,936 filed 11/17/2000, 09/714,936 filed 11/17/2000, 09/667,298 filed 9/22/00, 09/631,451 filed 8/3/00, and 09/598,042 filed 6/20/00, the disclosures of all of which are incorporated by reference herein in their entirety.

## FIELD OF THE INVENTION

The present invention relates to novel polynucleotides encoding proteins CG122, CG179, CG95, CG121, CG162, CG27, CG153, and CG168, which are related to proteins involved in lipid metabolism and cardiovascular disease, along with therapeutic, diagnostic and research utilities for these and related products.

## BACKGROUND

Lipoproteins are globular complexes made up of cholesteryl esters and/or triglycerides enveloped by amphiphilic phospholipids and apolipoproteins, that circulate in the bloodstream. The primary function of these molecules is to serve as carriers in the transport of nonpolar lipids. Lipoproteins are grouped into several classes based on their physical characteristics, and their associated lipids and apolipoprotein(s). The major classes include chylomicrons, chylomicron remnants, very low density lipoprotein (VLDL), intermediate density lipoprotein (IDL), low density lipoprotein (LDL), and high density lipoprotein (HDL). Chylomicrons contain apo AI, AII, CI, CII, CIII and E whereas chylomicron remnants are enriched for the B48 form of apo B, and apo E. VLDL contains the B100 form of apo B, apo CI, CII, CIII and E; IDL contains apo B100, CIII and E; LDL contains apo B100; and HDL contains apo AI and AII. Each of these major classes of lipoproteins also have sub-classes that contain different ratios of the primary apolipoproteins, and possibly other minor apolipoproteins. The primary function of chylomicrons and chylomicron remnants is to carry exogenous triglycerides and cholesteryl esters, whereas VLDL, IDL, LDL, and HDL, which differ in the ratio of

component triglycerides and cholesteryl esters, transport endogenous fats [Chappell et al. (1998) Prog lipid Res 37:393-422; Beiseigel (1998) Eur Heart J 19 Suppl.:A20-A23; Breslow (1993) Circ 87 Suppl. III:III-16-III-21].

Dietary fat enters circulation by incorporation into chylomicrons within the epithelial cells of the intestinal walls. The exogenous fat is then transported to skeletal and adipose tissue where the chylomicrons attach themselves to the capillary walls. Here, hydrolysis of chylomicrons into chylomicron remnants, mediated by lipoprotein lipase (LPL) or hepatic lipase (HL), releases fatty acids that are taken up by neighboring endothelial cells. Chylomicron remnants are removed from circulation, by internalization into the liver, through binding to the LDL receptor or LDL-receptor-related protein (also known as the a2-macroglobulin (LRP)). Binding of chylomicron remnants to the lipoprotein receptors is mediated by their associated apolipoproteins [Chappell et al. (1998) Prog Lipid Res 37:393-422; Beiseigel (1998) Eur Heart J 19 Suppl.:A20-A23; Breslow (1993) Circ 87 Suppl. III:III-16-III-21].

Endogenous triglycerides are synthesized in the liver and secreted into the plasma by incorporation into VLDL. VLDL is circulated to tissue capillaries where LPL and HL hydrolyze VLDL into VLDL remnants. VLDL remnants are cleared from the plasma by binding to LDL receptors and LRP in the liver via binding of apoE. However, most VLDL remnants undergo successive hydrolysis of their triglycerides, mediated by LPL and HL, into IDL and LDL such that the lipid portion of LDL is composed primarily of the remaining cholesteryl esters. LDL transports cholesteryl esters to a variety of cells including adrenal cortical cells, renal cells, hepatic cells, and lymphocytes. LDL is taken up by cells through binding to the LDL receptor and LRP via receptor-mediated endocytosis [Chappell et al. (1998) Prog lipid Res 37:393-422; Beiseigel (1998) Eur Heart J 19 Suppl.:A20-A23; Breslow (1993) Circ 87 Suppl. III:III-16-III-21]. Within these cells, the cholesteryl esters are delivered to the lysosome, where it is hydrolyzed into cholesterol by lysosomal acid lipase (LAL). In non-hepatic cells, cholesterol is used for membrane synthesis, hormone synthesis, and also in down-regulating LDL receptor synthesis. In the liver, cholesterol is either secreted into the bile or used to synthesize bile acids [Du et al. (1998) Mol Gen Meta 64:126-134].

HDL clears free cholesterol deposited, for example, as a by product of membrane turnover and/or cell death. In addition, HDL particles are primarily responsible for reverse cholesterol transport (RCT). RCT is a process in which excess cellular cholesterol is transported from peripheral tissues to the liver where it can be processed for excretion. The efflux of excess free cholesterol from peripheral cells is mediated primarily through the ATP-binding cassette transporter 1 (ABC1), also known as the cholesterol efflux regulatory protein (CERP) [Brooks-Wilson et al. (1999) Nat Genet 22:336-345]. The cholesterol is then taken up by Apo AI into HDL. Cholesterol carried by HDL is converted by lecithin-cholesterol acyltransferase (LCAT) into cholesteryl esters, which are then exchanged for triglycerides, present on VLDL and chylomicrons, by cholesteryl ester transfer protein (CETP) [Phillips et al. (1998) Atherosclerosis 137 Suppl:S13-S17; Stein et al. (1999) Atherosclerosis 44:285-301]. VLDL is then converted into cholesterol-rich LDL as discussed above. Thus, cholesterol is transported from extrahepatic cells to LDL via HDL, and LDL delivers cholesterol back to the liver. HDL is also taken up by the liver directly via component Apo E, and the LDL receptor and LRP mechanism described above [Beiseigel (1998) Eur Heart J Suppl A: A20-A23; Breslow (1993) Circ 87 suppl III:III-16-III-21; Chappell et al. (1998) Prog Lipid Res 37:393-422].

Lipoprotein composition and transport is regulated by apolipoproteins which serve as co-factors to enzymes involved in modifying lipoproteins, or as ligand recognition moieties for lipoprotein receptors. For example, apo CII acts as the co-factor for LPL, apo F regulates the activity of CETP, and apo E is important in receptor-mediated uptake of lipoproteins due to its high affinity for the LDL receptor and LRP [Chappell et al. (1998) Prog Lipid Res 37:393-422; Wang et al. (1999) J Biol Chem 274:1814-1820]. Lipid metabolism is also regulated by lipoprotein-processing proteins which include LPL, HL, LCAT, and CETP; and lipoprotein receptors such as LDL receptor, LRP, chylomicron remnant receptor, and scavenger receptors [Breslow (1993) Circ 87 suppl III:III-16-III-21; Hiltunen et al. (1998) Atherosclerosis S81-S88].

Abnormalities in lipid metabolism increase susceptibility to atherosclerosis and cardiovascular disease. Atherogenesis begins with lipid accumulation in the intima of the arterial wall due to retention of lipoproteins, such as LDL, by matrix proteoglycans. The phospholipids associated with LDL are hydrolzed by type II secretory non-pancreatic

phospholipase A2 (snpPLA$_2$) into free fatty acids (FFA) and lysophospholipids, both of which promote tissue inflammation [Hurt-Camejo et al. (1997) Atherosclerosis 132:1-8]. Cells present in the atherosclerotic lesions become activated leading to the production of inflammatory cytokines, snpPLA$_2$, LPL, macrophage colony stimulating factor (MCSF) and apo E, among others. These events result in changes in lipoprotein metabolism and the recruitment of macrophages to these sites. Both smooth muscle cells (SMC) and macrophages become lipid-filled cells, characteristic of atherosclerotic lesions, resulting from increased receptor-mediated uptake of modified LDL [Beiseigel (1998) Eur Heart J Suppl A: A20-A23; Hiltunen et a. (1998) Atherosclerosis S81-S88]. The signaling processes involved in a number of the processes described above involve receptor-activated cytosolic phospholipase C-β and A$_2$ [de Jonge et al. (1996) Mol Cell Biochem 157:199-210]. The resultant arterial plaques can become covered by fibrin clots and eventually occlude blood flow. Additionally, arterial plaques can rupture and break off the arterial wall. This can cause acute thrombotic events either at the site of rupture or as the circulating fragments block smaller vessels, and can lead to acute myocardial infarction, stroke, etc.

**SUMMARY OF THE INVENTION**

The compositions of the present invention include novel isolated polypeptides, in particular, novel human apolipoprotein, lipase, and lipoprotein receptor proteins and active variants thereof; isolated polynucleotides encoding such polypeptides, including recombinant DNA molecules, cloned genes; or degenerate variants thereof, especially naturally occurring variants such as allelic variants, antisense polynucleotide molecules, and antibodies that specifically recognize one or more epitopes present on such polypeptides, as well as hybridomas producing such antibodies.

The compositions of the present invention additionally include vectors, including expression vectors, containing the polynucleotides of the invention; cells genetically engineered to contain such polynucleotides; and cells genetically engineered to express such polynucleotides.

A nucleotide sequence encoding a protein designated CG122 (or C868) is set forth in SEQ ID NO: 1, and its deduced amino acid sequence is set forth in SEQ ID NO: 2. A

nucleotide sequence encoding a protein designated CG179 (or C355) is set forth in SEQ ID NO: 3, and its deduced amino acid sequence is set forth in SEQ ID NO: 4. An extended version of SEQ ID NO: 3 is set forth in SEQ ID NO: 16 and the deduced amino acid sequence is set forth in SEQ ID NO: 17. All of these proteins are believed to be new members of the apolipoprotein family. The polypeptide set out in SEQ ID NO: 2 is 366 amino acids in length, and amino acids 1-23 represent the putative signal peptide. eMatrix search results for SEQ ID NO: 2 showed an apolipoprotein plasma lipid transport domain (6.600e-14) at amino acids 75-130 and an apolipoprotein E precursor domain (4.779e-09) at amino acids 92-142; Pfam analysis showed an Apolipoprotein A1/A4/E family domain (1.6e-06) at amino acids 4 to 251. The polypeptides set out in SEQ ID NOS: 4 or 17 are 322 amino acids and 348 amino acids in length, respectively. eMatrix search results showed a phospholipase C signature (1.439e-20) at amino acids 295-314, an ICaBP type calcium binding protein signature (4.971e-09) at amino acids 135-172, and a Cyclin protein signature (5.114e-09) at amino acids 220-254 of SEQ ID NO: 17. CG122 shows 30% identity and 53% similarity at the amino acid level to pig apolipoprotein A-IV precursor protein (Genbank Accession No. AJ222966), 28% identity and 49% similarity to apolipoprotein A-IV precursor protein from macaque (Genbank Accession No. X68361), and 27% identity and 48% similarity to chick apolipoprotein A-IV (Genbank Accession No. Y16534). CG122 shows 100% identity at the amino acid level to a sequence of GENBANK Accession No. gi12406730 and a sequence from Int'l Publication No. WO20/037491. CG179 shows 59% identity and 76% similarity at the amino acid level to human TNF-inducible protein (Genbank Accession No. AF070675), 40% identity and 59% similarity to human protein dJ6802.1 (Genbank Accession No.Z82215), and 39% identity and 58% similarity to human apolipoprotein L precursor (Genbank Accession No. AF019225). Figure 1 shows an alignment of CG179 (C355) SEQ ID NOS: 4, 17, a sequence of GENBANK Accession No. gi12408272, and a sequence from PCT Publication No. WO99/31236, and shows that amino acids 1-93 of SEQ ID NO: 4 or 17 are missing from that published sequence. Thus, a preferred CG179 polypeptide comprises one or more (or preferably 10 or more) of amino acids 1-93 of SEQ ID NO: 4 or 17. Additional family members can be identified using either SEQ ID NO: 1 or SEQ ID NO: 3 or fragments thereof as a molecular probe.

A nucleotide sequence encoding a lipase protein designated CG95 (or C870) is set forth in SEQ ID NO: 5, and its deduced amino acid sequence is set forth in SEQ ID NO: 6. Analysis of the amino acid sequence reveals possible proteolytic cleavage sites at either amino acid residue 21 or 24 of SEQ ID NO: 6. As a result, either amino acids 1-24 or more likely amino acids 1-21 are predicted to be a signal peptide. Therefore, either amino acids 22-145 or amino acids 25-145 comprise a secreted, mature protein with lipase function. eMatrix search results on SEQ ID NO: 6 showed phospholipase A2 signatures at amino acids 44-72, 56-75, 37-56, 104-121, 104-120, 79-98; Pfam search results showed phospholipase A2 domains (1.1e-47) at amino acids 21 to 145. A nucleotide sequence encoding a lipase protein designated CG121 (or C592) is set forth in SEQ ID NO: 7, and its deduced amino acid sequence is set forth in SEQ ID NO: 8. A slightly different and shorter version of SEQ ID NO: 7 is set forth in SEQ ID NO: 18 and the deduced amino acid sequence is set forth in SEQ ID NO: 19. A nucleotide sequence encoding a lipase protein designated CG162 (or C59) is set forth in SEQ ID NO: 9. One of skill in the art could determine the corresponding amino acid sequence using techniques well known in the art to translate and analyze all possible six frames. The present invention contemplates proteins encoded by each of the six possible reading frames, in particular those proteins, polypeptides or fragments thereof exhibiting homology to lysosomal acid lipases are preferred. An extended version of SEQ ID NO: 9 is set forth in SEQ ID NO: 20 and the deduced amino acid sequence is set forth in SEQ ID NO: 21. CG95 and CG121 are believed to be new members of the phospholipase family. CG162 is believed to be a novel lysosomal acid lipase. The polypeptide set out in SEQ ID NO: 6 is 145 amino acids in length. The polypeptides set out in SEQ ID NOS: 8 or 19 are 567 amino acids or 340 amino acids in length, respectively. Pfam analysis of SEQ ID NO: 19 showed a Phosphatidylinositol-specific phospholipase domain (5.6e-16) at amino acids 291 to 326 and a PH domain (phospholipid binding) (1.8e-11) at amino acids 17 to 124; an alpha/beta hydrolase fold (8.9e-13) was also predicted at amino acids 111 to 390. The polypeptide set out in SEQ ID NO: 21 is 409 amino acids in length, and amino acids 1-19 represent the putative signal peptide. The polypeptides of SEQ ID NO: 6 and SEQ ID NO: 8 display amino acid homology with the human $PLA_2$ and PLC respectively. CG95 shows 47% identity and 63% similarity at the amino acid level to rat phospholipase A2 (Genbank Accession Nos. X51529 and M37127), 47% identity and

63% similarity to rat phospholipase A2 membrane associated precursor (Genbank Accession No. D00523), and 47% identity and 63% similarity to human synovial phospholipase A2 (Genbank Accession Nos. M22431 and M22430). CG95 shows nearly 100% identity at the amino acid level to a sequence of GENBANK Accession No. gi5771420 and a sequence from Int'l Publication No. WO 20/024911. CG121 shows 73% identity and 82% similarity at the amino acid level to bovine 1-phosphotidylinositol-4,5-bisphosphate phosphodiesterase delta-2 (Genbank Accession No. S14113), 65% identity and 76% similarity to rat phospholipase C delta-4 (Genbank Accession No. U16655), and 65% identity and 76% similarity to rat phospholipase C delta-4 (Genbank Accession No. D50455). Figure 2 shows an alignment of CG121(C592) SEQ ID NOS: 8, 19, a sequence of GENBANK Accession No. gi1304189, and a sequence from GENBANK Accession No. gi571466, and shows that amino acids 326-340 of SEQ ID NO: 8 or 19 are missing from that published sequence. Thus, a preferred CG179 polypeptide comprises one or more (or preferably 10 or more) of amino acids 326-340 of SEQ ID NO: 8 or 19. Additional family members can be identified using either SEQ ID NO: 5 or SEQ ID NO: 7 as a molecular probe. CG162 shows 60% identity and 75% similarity at the amino acid level to human lysosomal acid lipase (Genbank Accession No. U04285), 60% identity and 75% similarity to human lysosomal acid lipase (Genbank Accession No. U08464), and 63% identity and 78% similarity to human lysosomal acid lipase precursor (Genbank Accession No. M74775). Figure 3 shows an alignment of SEQ ID NO: 21, a sequence from GENBANK Accession No. gi434306, and a sequence from Int'l Publication No. WO 86/03778 and shows that SEQ ID NO: 21 exhibits about 60% and 52% identity to these proteins, identified putatively as a sterol esterase and pregastric lipase, respectively. Additional family members can be identified using SEQ ID NO: 9 as a molecular probe.

A nucleotide sequence encoding a receptor protein designated CG27 (or C869) is set forth in SEQ ID NO: 10, and its deduced amino acid sequence is set forth in SEQ ID NO: 11. Four additional variant nucleotide sequences are set forth in SEQ ID NOS: 22, 24, 26 and 44 and their respective deduced amino acid sequences are set forth in SEQ ID NOS: 23, 25, 27 and 45. A nucleotide sequence encoding a receptor protein designated CG153 (or C593) is set forth in SEQ ID NO: 12, and its deduced amino acid sequence is set forth in SEQ ID NO: 13. Two additional variant nucleotide sequences are set forth in

SEQ ID NOS: 28 and 30, and their respective deduced amino acid sequences are set forth in SEQ ID NOS: 29 and 31. A nucleotide sequence encoding a receptor protein designated CG 168 (or C595) is set forth in SEQ ID NO: 14, and its deduced amino acid sequence is set forth in SEQ ID NO: 15. SEQ ID NO: 14 contains two possible start codons, one at nucleotide position 149 and a second possible start codon at nucleotide position 260. One of skill in the art using well known techniques, *i.e.*, using primer extension, can determine the correct start codon. An extended version of SEQ ID NO: 14 is set forth in SEQ ID NO: 32 and the deduced amino acid sequence is set forth in SEQ ID NO: 33. The polypeptides set out in SEQ ID NOS: 11, 23, 25 or 27 are 288, 280, 314 or 247 amino acids in length, respectively. eMatrix search results showed a C-type lectin domain (2.080e-11) at amino acids 148-166 of SEQ ID NO: 23, amino acids 175-193 of SEQ ID NO: 25, and amino acids 115-133 of SEQ ID NO: 27; Pfam search results also showed a Lectin C-type domain (5.1e-05) at amino acids 163 to 257 of SEQ ID NO: 23, amino acids 190 to 284 of SEQ ID NO: 25, and amino acids 130 to 224 of SEQ ID NO: 27. The polypeptides set out in SEQ ID NO: 13, 29 or 31 are 732 amino acids, 753 amino acids or 608 amino acids in length, respectively, and amino acids 1-25 represent the putative signal peptide in all of these polypeptides. eMatrix search results for SEQ ID NO: 29 showed a Speract receptor repeat proteins domain (6.250e-27) at amino acids 311-366, a lysyl oxidase signature (1.522e-25) at amino acids 675-704 and a lysyl oxidase copper-binding region signature (5.500e-25) at amino acids 652-692, a Speract receptor repeat proteins domain (5.442e-24) at amino acids 49-104, a lysyl oxidase copper-binding region (5.671e-24) at amino acids 584-621,a lysyl oxidase signature (4.667e-20) at amino acids 589-618, a lysyl oxidase signature (4.000e-16) at amino acids 617-645, a lysyl oxidase copper-binding region (7.257e-15) at amino acids 692-733 a lysyl oxidase copper-binding region (8.327e-14) at amino acids 538-585, a lysyl oxidase copper-binding region (2.102e-13) at amino acids 620-651, a lysyl oxidase signature (5.500e-13) at amino acids 704-732, a Speract receptor repeat proteins domain (7.840e-13) at amino acids 134-145, a Speract receptor repeat proteins domain (3.972e-12) at amino acids 180-235, a speract receptor signature (5.721e-11) at amino acids 417-434, a speract receptor signature (7.000e-11) at amino acids 395-408, a Speract receptor repeat protein domain (8.017e-11) at amino acids 396-407, a speract receptor signature (9.250e-11) at amino acids 133-146, a speract receptor signature (2.469e-10) at

amino acids 341-352, a lysyl oxidase signature (2.514e-10) at amino acids 533-562, a speract receptor signature (2.746e-10) at amino acids 307-324, a Speract receptor repeat proteins domain (3.526e-10) at amino acids 425-480, a speract receptor signature (4.724e-10) at amino acids 372-387, a speract receptor signature (6.311e-10) at amino acids 64-76, and a speract receptor signature (7.429e-09) at amino acids 488-503; Pfam analysis of SEQ ID NO: 29 also showed a Lysyl oxidase domain (2.9e-173) at amino acids 529 to 732 and Scavenger receptor cysteine-rich domains (7e-82) at amino acids 51 to 145, 183 to 282, 310 to 407 and 420 to 525. Pfam analysis of SEQ ID NO: 31 showed Scavenger receptor cysteine-rich domains at amino acids 51 to 145, 165 to 262, and 275 to 380 and a lysyl oxidase domain at amino acids 384 to 587. The polypeptides set out in SEQ ID NO: 15 or 33 are 639 amino acids or 4636 amino acids in length, respectively. eMatrix and Pfam analysis of SEQ ID NO: 33 show over 100 LDL receptor signature repeats as well as numerous EGF-like domains. CG27 and CG168 are believed to be new members of the LDL receptor family. CG27 shows 31% identity and 51% similarity at the amino acid level to bovine lectin-like oxidized LDL receptor (Genbank Accession No. D89049), 29% identity and 48% similarity to human oxidized low density lipoprotein (lectin-like) receptor (Genbank Accession Nos. AB010710, AF035776, and AF079167), and 28% identity and 50% similarity to rat endothelial receptor for oxidized low density lipoprotein (Genbank Accession No. AB0005900). Figure 4 shows an alignment of CG27 (C869) SEQ ID NOS: 11, 23, 25, 27, a sequence of GENBANK Accession No. gi7110216, and a sequence from Int'l Publication No. WO 99/13066, and shows that amino acids 111-138 of SEQ ID NO: 11 and 25, corresponding to exon 2, are missing from that published sequence. Thus, a preferred CG27 polypeptide comprises one or more (or preferably 10 or more) of amino acids 111-138 of SEQ ID NO: 11 or 25. CG168 shows 59% identity and 74% similarity at the amino acid level to chick alpha-2-macroglobulin receptor precursor (Genbank Accession No. X74904), 58% identity and 74% similarity to murine AM2 receptor (Genbank Accession No. X67469), and 58% identity and 73% similarity to human low density lipoprotein-related protein 1 (alpha-2-macroglobulin receptor) (Genbank Accession No. X13916). Figure 5 shows an alignment of CG168 (C595) SEQ ID NOS: 15 and 33 with a sequence that may be disclosed in Liu et al., Cancer Res. 60(7):1961-1967 (2000), and shows that amino acids 1-37 are missing from that sequence. Additional family members can be identified using

SEQ ID NO: 10 or 14 as a molecular probe. CG153 shows 90% identity and 93% similarity at the amino acid level to murine lysyl oxidase-related protein 2 (Genbank Accession No. AF053368), and 54% identity and 71% similarity to human lysyl oxidase-related protein 2 (Genbank Accession No. U89942). Figure 6 shows an alignment of CG153 (C593) SEQ ID NOS: 13, 29, 31, a sequence of GENBANK Accession No. gi3978171, and a sequence from Int'l Publication No. WO 20/0044910. Additional family members can be identified using SEQ ID NO: 12 as a molecular probe.

The polynucleotides of the invention include naturally occurring or wholly or partially synthetic DNA, e.g., cDNA and genomic DNA, and RNA, e.g., mRNA. The isolated polynucleotides of the invention include, but are not limited to, a polynucleotide encoding a polypeptide comprising the amino acid sequence of SEQ ID NOS: 2, 4, 6, 8, 11, 13, 15, 17, 19, 21, 23, 25, 27, 29, 31, 33, 35, 37, 39, 41, 43 or 45; fragments thereof or the corresponding full length or mature proteins. The mature portion of each protein can be determined by expression of the corresponding cDNA in an appropriate host cell. The isolated polynucleotides of the invention further include, but are not limited to, a polynucleotide comprising the nucleotide sequence of SEQ ID NO: 1, 3, 5, 7, 9, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, 34, 36, 38, 40, 42 or 44; a polynucleotide comprising the full length protein coding sequence of SEQ ID NO: 1, 3, 5, 7, 9, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, 34, 36, 38, 40, 42 or 44; and a polynucleotide comprising the nucleotide sequence of the mature protein coding sequence of SEQ ID NO: 1, 3, 5, 7, 9, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, 34, 36, 38, 40, 42 or 44. The polynucleotides of the present invention also include, but are not limited to, polynucleotides that encode polypeptides with biological activity, that hybridize under stringent hybridization conditions to the complement of (a) the nucleotide sequence of SEQ ID NO: 1, 3, 5, 7, 9, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, 34, 36, 38, 40, 42 or 44; or (b) a nucleotide sequence encoding the amino acid sequence of SEQ ID NO: 2, 4, 6, 8, 11, 13, 15, 17, 19, 21, 23, 25, 27, 29, 31, 33, 35, 37, 39, 41, 43 or 45; or a polynucleotide which is an allelic variant of any polynucleotide recited above; a polynucleotide which encodes a species homolog (e.g. orthologs) of any of the proteins recited above; or a polynucleotide that encodes a polypeptide comprising a specific domain or truncation of the polypeptide having an amino acid sequence of SEQ ID NO: 2, 4, 6, 8, 11, 13, 15, 17, 19, 21, 23, 25, 27, 29, 31, 33, 35, 37, 39, 41, 43 or 45. The

polynucleotides of the invention additionally include the complement of any of the polynucleotides recited above.

The isolated polypeptides of the invention include, but are not limited to, a polypeptide comprising the amino acid sequence of SEQ ID NO: 2, 4, 6, 8, 11, 13, 15, 17, 19, 21, 23, 25, 27, 29, 31, 33, 35, 37, 39, 41, 43 or 45; fragments thereof or the corresponding full length or mature protein. Polypeptides of the invention also include polypeptides with biological activity that are encoded by (a) polynucleotides set out in SEQ ID NO: 1, 3, 5, 7, 9, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, 34, 36, 38, 40, 42 or 44; or (b) polynucleotides that hybridize to the complement of the polynucleotides of (a) under stringent hybridization conditions. Biologically or immunologically active variants of the protein sequence of SEQ ID NO: 2, 4, 6, 8, 11, 13, 15, 17, 19, 21, 23, 25, 27, 29, 31, 33, 35, 37, 39, 41, 43 or 45; and "substantial equivalents" thereof (e.g., with 65%, 70%, 75%, 80%, 85%, 90%, 95%, 98% or 99% amino acid sequence identity) that preferably retain biological activity are also contemplated. The polypeptides of the invention may be wholly or partially chemically synthesized but are preferably produced by recombinant means using the genetically engineered cells (e.g. host cells) of the invention.

Protein compositions of the present invention may further comprise an acceptable carrier, such as a hydrophilic, e.g., pharmaceutically acceptable, carrier.

The invention also relates to methods for producing polypeptides of the invention comprising growing a culture of the cells of the invention in a suitable culture medium under conditions permitting expression of the desired polypeptide, and purifying the protein from the cells or the culture medium in which the cells are grown. Preferred embodiments include those in which the protein produced by such process is a mature form of the protein.

Polynucleotides according to the invention have numerous applications in a variety of techniques known to those skilled in the art of molecular biology. These techniques include use as hybridization probes, use as oligomers for PCR, use for chromosome and gene mapping, use in the recombinant production of protein, and use in generation of antisense DNA or RNA, their chemical analogs and the like. For example, when the expression of an mRNA is largely restricted to a particular cell or tissue type, polynucleotides of the invention can be used as hybridization probes to detect or quantify

the presence of the particular cell or tissue mRNA in a sample using, e.g., *in situ* hybridization.

In other exemplary embodiments, the polynucleotides are used in diagnostics as expressed sequence tags for identifying expressed genes or, as well known in the art and exemplified by Vollrath et al., Science 258:52-59 (1992), as expressed sequence tags for physical mapping of the human genome.

The polypeptides according to the invention can be used in a variety of conventional procedures and methods that are currently applied to other proteins. For example, a polypeptide of the invention can be used to generate an antibody that specifically binds the polypeptide. Such antibodies, particularly monoclonal antibodies, are useful for detecting or quantitating the polypeptide in tissue. The polypeptides of the invention can also be used as molecular weight markers, and as a food supplement.

Methods are also provided for preventing, treating, or ameliorating a medical condition which comprises the step of administering to a mammalian subject a therapeutically effective amount of a composition comprising a protein or polypeptide of the present invention and a pharmaceutically acceptable carrier.

In particular, the polypeptides and polynucleotides of the invention may play a role in disorders involving lipid metabolism, thrombosis, and cardiovascular disease, including occlusive cardiovascular diseases such as myocardial infarction, cerebral ischemia, and angina; arterial thrombosis, such as coronary artery thrombosis and resulting myocardial infarction; cerebral artery thrombosis or intracardiac thrombosis (due to, *e.g.*, atrial fibrillation) and resulting stroke, and other peripheral arterial thrombosis and occlusion; conditions associated with venous thrombosis, such as deep venous thrombosis and pulmonary embolism; conditions associated with exposure of the patient's blood to a foreign or injured tissue surface, including diseased heart valves, mechanical heart valves, vascular grafts, and other extracorporeal devices such as intravascular cannulas, vascular access shunts in hemodialysis patients, hemodialysis machines and cardiopulmonary bypass machines; and conditions associated with coagulapathies, such as hypercoagulability and disseminated intravascular coagulopathy.

The methods of the present invention further relate to methods for detecting the presence of the polynucleotides or polypeptides of the invention in a sample. Such methods can, for example, be utilized as part of prognostic and diagnostic evaluation of

disorders as recited herein and for the identification of subjects exhibiting a predisposition to such conditions. The invention also provides kits comprising polynucleotide probes and/or monoclonal antibodies, and optionally quantitative standards, for carrying out methods of the invention. Furthermore, the invention provides methods for evaluating the efficacy of drugs, and monitoring the progress of patients, involved in clinical trials for the treatment of disorders as recited herein.

The invention also provides methods for the identification of compounds that modulate (i.e., increase or decrease) the expression or activity of the polynucleotides and/or polypeptides of the invention. Such methods can be utilized, for example, for the identification of compounds that can ameliorate symptoms of disorders as recited herein. Such methods can include, but are not limited to, assays for identifying compounds and other substances that interact with (e.g., bind to) the polypeptides of the invention.

The methods of the invention also include methods for the treatment of disorders as recited herein which may involve the administration of the polynucleotides or polypeptides of the invention to individuals exhibiting symptoms or tendencies related to disorders as recited herein. In addition, the invention encompasses methods for treating diseases or disorders as recited herein comprising the step of administering compounds and other substances that modulate the overall activity of the target CG122, CG179, CG 95, CG121, CG162, CG27, CG153, and CG168 gene products. Compounds and other substances can effect such modulation either on the level of target gene/protein expression or target protein activity.

## BRIEF DESCRIPTION OF THE DRAWINGS

Figures 1A-1B show an alignment of CG179 (C355) SEQ ID NOS: 4, 17, a sequence of GENBANK Accession No. gi12408272, and a sequence from PCT Publication No. WO99/31236.

Figures 2A-2D shows an alignment of CG121(C592) SEQ ID NOS: 8, 19, a sequence of GENBANK Accession No. gi1304189, and a sequence from GENBANK Accession No. gi571466.

Figures 3A-3B shows an alignment of SEQ ID NO: 21, a sequence from GENBANK Accession No. gi434306, and a sequence from Int'l Publication No. WO

86/03778 and shows that SEQ ID NO: 21 exhibits about 60% and 52% identity to these proteins, identified putatively as a sterol esterase and pregastric lipase, respectively.

Figures 4A-4B shows an alignment of CG27 (C869) SEQ ID NOS: 11, 23, 25, 27, a sequence of GENBANK Accession No. gi7110216, and a sequence from Int'l Publication No. WO 99/13066.

Figures 5A-5P shows an alignment of CG168 (C595) SEQ ID NOS: 15 and 33 with a sequence that may be disclosed in Liu et al., Cancer Res. 60(7):1961-1967 (2000).

Figures 6A-6D shows an alignment of CG153 (C593) SEQ ID NOS: 13, 29, 31, a sequence of GENBANK Accession No. gi3978171, and a sequence from Int'l Publication No. WO 20/0044910.

## DETAILED DESCRIPTION OF THE INVENTION

### 1. DEFINITIONS

The term "nucleotide sequence" refers to a heteropolymer of nucleotides or the sequence of these nucleotides. The terms "nucleic acid" and "polynucleotide" are also used interchangeably herein to refer to a heteropolymer of nucleotides. Generally, nucleic acid segments provided by this invention may be assembled from fragments of the genome and short oligonucleotide linkers, or from a series of oligonucleotides, or from individual nucleotides, to provide a synthetic nucleic acid which is capable of being expressed in a recombinant transcriptional unit comprising regulatory elements derived from a microbial or viral operon, or a eukaryotic gene.

The terms "oligonucleotide fragment" or a "polynucleotide fragment", "portion," or "segment" is a stretch of polypeptide nucleotide residues which is long enough to use in polymerase chain reaction (PCR) or various hybridization procedures to identify or amplify identical or related parts of mRNA or DNA molecules.

The terms "oligonucleotides" or "nucleic acid probes" are prepared based on the polynucleotide sequences provided in the present invention. Oligonucleotides comprise portions of such a polynucleotide sequence having at least about 15 nucleotides and usually at least about 20 nucleotides. Nucleic acid probes comprise portions of such a polynucleotide sequence having fewer nucleotides than about 6 kb, usually fewer than about 1 kb. After appropriate testing to eliminate false positives, these probes may, for example, be used to determine whether specific mRNA molecules are present in a cell or

tissue or to isolate similar nucleic acid sequences from chromosomal DNA as described by Walsh et al. (Walsh, P.S. et al., 1992, PCR Methods Appl 1:241-250).

The term "probes" includes naturally occurring or recombinant or chemically synthesized single- or double-stranded nucleic acids. They may be labeled by nick translation, Klenow fill-in reaction, PCR, or other methods well known in the art. Probes of the present invention, their preparation and/or labeling are elaborated in Sambrook, J. et al., 1989, Molecular Cloning: A Laboratory Manual, Cold Spring Harbor Laboratory, NY; or Ausubel, F.M. et al., 1989, Current Protocols in Molecular Biology, John Wiley & Sons, New York NY, both of which are incorporated herein by reference in their entirety.

The term "stringent" is used to refer to conditions that are commonly understood in the art as stringent. Stringent conditions can include highly stringent conditions (e.g., hybridization to filter-bound DNA in 0.5 M $NaHPO_4$, 7% sodium dodecyl sulfate (SDS), 1 mM EDTA at 65°C, and washing in 0.1X SSC/0.1% SDS at 68°C), and moderately stringent conditions (e.g., washing in 0.2X SSC/0.1% SDS at 42°C). Other exemplary hybridization conditions are described herein in the examples.

In instances wherein hybridization of deoxyoligonucleotides is concerned, additional exemplary stringent hybridization conditions include washing in 6X SSC/0.05% sodium pyrophosphate at 37°C (for 14-base oligos), 48°C (for 17-base oligos), 55°C (for 20-base oligos), and 60°C (for 23-base oligos).

The term "recombinant," when used herein to refer to a polypeptide or protein, means that a polypeptide or protein is derived from recombinant (e.g., microbial, insect, or mammalian) expression systems. "Microbial" refers to recombinant polypeptides or proteins made in bacterial or fungal (e.g., yeast) expression systems. As a product, "recombinant microbial" defines a polypeptide or protein essentially free of native endogenous substances and unaccompanied by associated native glycosylation. Polypeptides or proteins expressed in most bacterial cultures, e.g., *E. coli*, will be free of glycosylation modifications; polypeptides or proteins expressed in yeast will have a glycosylation pattern in general different from those expressed in mammalian cells.

The term "recombinant expression vehicle or vector" refers to a plasmid or phage or virus or vector, for expressing a polypeptide from a DNA (RNA) sequence. An expression vehicle can comprise a transcriptional unit comprising an assembly of (1) a genetic element or elements having a regulatory role in gene expression, for example,

promoters or enhancers, (2) a structural or coding sequence which is transcribed into mRNA and translated into protein, and (3) appropriate transcription initiation and termination sequences. Structural units intended for use in yeast or eukaryotic expression systems preferably include a leader sequence enabling extracellular secretion of translated protein by a host cell. Alternatively, where recombinant protein is expressed without a leader or transport sequence, it may include an N-terminal methionine residue. This residue may or may not be subsequently cleaved from the expressed recombinant protein to provide a final product.

The term "recombinant expression system" means host cells which have stably integrated a recombinant transcriptional unit into chromosomal DNA or carry the recombinant transcriptional unit extrachromosomally. Recombinant expression systems as defined herein will express heterologous polypeptides or proteins upon induction of the regulatory elements linked to the DNA segment or synthetic gene to be expressed. This term also means host cells which have stably integrated a recombinant genetic element or elements having a regulatory role in gene expression, for example, promoters or enhancers. Recombinant expression systems as defined herein will express polypeptides or proteins endogenous to the cell upon induction of the regulatory elements linked to the endogenous DNA segment or gene to be expressed. The cells can be prokaryotic or eukaryotic.

The term "open reading frame," ORF, means a series of nucleotide triplets coding for amino acids without any termination codons and is a sequence translatable into protein.

The term "expression modulating fragment," EMF, means a series of nucleotides which modulates the expression of an operably linked ORF or another EMF.

As used herein, a sequence is said to "modulate the expression of an operably linked sequence" when the expression of the sequence is altered by the presence of the EMF. EMFs include, but are not limited to, promoters, and promoter modulating sequences (inducible elements). One class of EMFs are fragments which induce the expression or an operably linked ORF in response to a specific regulatory factor or physiological event.

As used herein, an "uptake modulating fragment," UMF, means a series of nucleotides which mediate the uptake of a linked DNA fragment into a cell. UMFs can

be readily identified using known UMFs as a target sequence or target motif with the computer-based systems described below.

The presence and activity of a UMF can be confirmed by attaching the suspected UMF to a marker sequence. The resulting nucleic acid molecule is then incubated with an appropriate host under appropriate conditions and the uptake of the marker sequence is determined. As described above, a UMF will increase the frequency of uptake of a linked marker sequence.

The term "active" refers to those forms of the polypeptide which retain the biologic and/or immunologic activities of any naturally occurring polypeptide. According to the invention, the term “biologically active” with reference to the apolipoprotein-like polypeptides of the invention means that the polypeptide retains at least one of the biological activities of CG122 or CG179, preferably the apolipoprotein activity. The term “biologically active” with reference to the lipase-like polypeptides of the invention means that the polypeptide retains at least one of the biological activities of CG95, CG121, or CG162, preferably the lipase activity. The term “biologically active” with reference to the lipoprotein receptor-like polypeptides of the invention means that the polypeptide retain at least one of the biological activities of CG27, CG153, or CG168, preferably lipoprotein receptor activity. The term “immunologically active” with reference to the apolipoprotein, lipase, or lipoprotein receptor polypeptides of the invention means that the polypeptide retains at least one of the immunologic or antigenic activities of CG122, CG179, CG95, CG121, CG162, CG27, CG153 or CG168.

The term "naturally occurring polypeptide" refers to polypeptides produced by cells that have not been genetically engineered and specifically contemplates various polypeptides arising from post-translational modifications of the polypeptide including, but not limited to, acetylation, carboxylation, glycosylation, phosphorylation, lipidation and acylation.

The term "derivative" refers to polypeptides chemically modified by such techniques as ubiquitination, labeling (e.g., with radionuclides or various enzymes), pegylation (derivatization with polyethylene glycol) and insertion or substitution by chemical synthesis of amino acids such as ornithine, which do not normally occur in human proteins.

The term "variant" (or “analog”) refers to any polypeptide differing from naturally occurring polypeptides by amino acid insertions, deletions, and substitutions, created using, for example, recombinant DNA techniques. Guidance in determining which amino acid residues may be replaced, added or deleted without abolishing activities of interest, such as apolipoprotein, lipase, or lipoprotein receptor activity, may be found by comparing the sequence of the particular polypeptide with that of homologous human or other mammalian apolipoprotein, lipase, or lipoprotein receptor polypeptides and minimizing the number of amino acid sequence changes made in regions of high homology (conserved regions) or by replacing amino acids with consensus sequence.

Preferably, amino acid "substitutions" are the result of replacing one amino acid with another amino acid having similar structural and/or chemical properties, i.e., conservative amino acid replacements. “Conservative” amino acid substitutions may be made on the basis of similarity in polarity, charge, solubility, hydrophobicity, hydrophilicity, and/or the amphipathic nature of the residues involved. For example, nonpolar (hydrophobic) amino acids include alanine, leucine, isoleucine, valine, proline, phenylalanine, tryptophan, and methionine; polar neutral amino acids include glycine, serine, threonine, cysteine, tyrosine, asparagine, and glutamine; positively charged (basic) amino acids include arginine, lysine, and histidine; and negatively charged (acidic) amino acids include aspartic acid and glutamic acid. "Insertions" or "deletions" are typically in the range of about 1 to 5 amino acids. The variation allowed may be experimentally determined by systematically making insertions, deletions, or substitutions of amino acids in a polypeptide molecule using recombinant DNA techniques and assaying the resulting recombinant variants for activity.

Alternatively, where alteration of function is desired, insertions, deletions or non-conservative alterations can be engineered to produce altered polypeptides. Such alterations can, for example, alter one or more of the biological functions or biochemical characteristics of the polypeptides of the invention. For example, such alterations may change polypeptide characteristics such as ligand-binding affinities, interchain affinities, or degradation/turnover rate. Further, such alterations can be selected so as to generate polypeptides that are better suited for expression, scale up and the like in the host cells chosen for expression. For example, cysteine residues can be deleted or substituted with another amino acid residue in order to eliminate disulfide bridges.

As used herein, "substantially equivalent" can refer both to nucleotide and amino acid sequences, for example a mutant sequence, that varies from a reference sequence by one or more substitutions, deletions, or additions, the net effect of which does not result in an adverse functional dissimilarity between the reference and subject sequences. Typically, such a substantially equivalent sequence varies from one of those listed herein by no more than about 20% (i.e., the number of individual residue substitutions, additions, and/or deletions in a substantially equivalent sequence, as compared to the corresponding reference sequence, divided by the total number of residues in the substantially equivalent sequence is about 0.2 or less). Such a sequence is said to have 80% sequence identity to the listed sequence. In one embodiment, a substantially equivalent, e.g., mutant, sequence of the invention varies from a listed sequence by no more than 10% (90% sequence identity); in a variation of this embodiment, by no more than 5% (95% sequence identity); and in a further variation of this embodiment, by no more than 2% (98% sequence identity). Substantially equivalent, e.g., mutant, amino acid sequences according to the invention generally have at least 95% sequence identity with a listed amino acid sequence, whereas substantially equivalent nucleotide sequence of the invention can have lower percent sequence identities, taking into account, for example, the redundancy or degeneracy of the genetic code. For the purposes of the present invention, sequences having substantially equivalent biological activity and substantially equivalent expression characteristics are considered substantially equivalent. For the purposes of determining equivalence, truncation of the mature sequence (e.g., via a mutation which creates a spurious stop codon) should be disregarded. Sequence identity may be determined, e.g., using the Jotun Hein method.

Nucleic acid sequences encoding such substantially equivalent sequences, e.g., sequences of the recited percent identities, can routinely be isolated and identified via standard hybridization procedures well known to those of skill in the art.

Where desired, an expression vector may be designed to contain a "signal or leader sequence" which will direct the polypeptide through the membrane of a cell. Such a sequence may be naturally present on the polypeptides of the present invention or provided from heterologous protein sources by recombinant DNA techniques.

A polypeptide "fragment," "portion," or "segment" is a stretch of amino acid residues of at least about 5 amino acids, often at least about 7 amino acids, typically at

least about 9 to 13 amino acids, and, in various embodiments, at least about 17 or more amino acids. To be active, any polypeptide must have sufficient length to display biologic and/or immunologic activity.

Alternatively, recombinant variants encoding these same or similar polypeptides may be synthesized or selected by making use of the "redundancy" in the genetic code. Various codon substitutions, such as the silent changes which produce various restriction sites, may be introduced to optimize cloning into a plasmid or viral vector or expression in a particular prokaryotic or eukaryotic system. Mutations in the polynucleotide sequence may be reflected in the polypeptide or domains of other peptides added to the polypeptide to modify the properties of any part of the polypeptide, to change characteristics such as ligand-binding affinities, interchain affinities, or degradation/turnover rate.

The term "activated" cells as used in this application are those which are engaged in extracellular or intracellular membrane trafficking, including the export of neurosecretory or enzymatic molecules as part of a normal or disease process.

The term "purified" as used herein denotes that the indicated nucleic acid or polypeptide is present in the substantial absence of other biological macromolecules, e.g., polynucleotides, proteins, and the like. In one embodiment, the polynucleotide or polypeptide is purified such that it constitutes at least 95% by weight, more preferably at least 99.8% by weight, of the indicated biological macromolecules present (but water, buffers, and other small molecules, especially molecules having a molecular weight of less than 1000 daltons, can be present).

The term "isolated" as used herein refers to a nucleic acid or polypeptide separated from at least one other component (e.g., nucleic acid or polypeptide) present with the nucleic acid or polypeptide in its natural source. In one embodiment, the nucleic acid or polypeptide is found in the presence of (if anything) only a solvent, buffer, ion, or other component normally present in a solution of the same. The terms "isolated" and "purified" do not encompass nucleic acids or polypeptides present in their natural source.

The term "infection" refers to the introduction of nucleic acids into a suitable host cell by use of a virus or viral vector.

The term "transformation" means introducing DNA into a suitable host cell so that the DNA is replicable, either as an extrachromosomal element, or by chromosomal integration.

The term "transfection" refers to the taking up of an expression vector by a suitable host cell, whether or not any coding sequences are in fact expressed.

The term "intermediate fragment" means a nucleic acid between 5 and 1000 bases in length, and preferably between 10 and 40 bp in length.

The term "secreted" includes a protein that is transported across or through a membrane, including transport as a result of signal sequences in its amino acid sequence when it is expressed in a suitable host cell. "Secreted" proteins include without limitation proteins secreted wholly (e.g., soluble proteins) or partially (e.g., receptors) from the cell in which they are expressed. "Secreted" proteins also include without limitation proteins which are transported across the membrane of the endoplasmic reticulum. "Secreted" proteins are also intended to include proteins containing non-typical signal sequences (e.g. Interleukin-1 Beta, see Krasney, P.A. and Young, P.R. (1992) Cytokine 4(2): 134 -143) and factors released from damaged cells (e.g. Interleukin-1 Receptor Antagonist, see Arend, W.P. et. al. (1998) Annu. Rev. Immunol. 16:27-55)

Each of the above terms is meant to encompasses all that is described for each, unless the context dictates otherwise.

## NUCLEIC ACIDS AND POLYPEPTIDES OF THE INVENTION

Nucleotide and amino acid sequences of the invention are reported below. Fragments of the proteins of the present invention which are capable of exhibiting biological activity are also encompassed by the present invention. Fragments of the proteins may be in linear form or they may be cyclized using known methods, for example, as described in H. U. Saragovi, et al., Bio/Technology 10, 773-778 (1992) and in R. S. McDowell, et al., J. Amer. Chem. Soc. 114, 9245-9253 (1992), both of which are incorporated herein by reference. Such fragments may be fused to carrier molecules such as immunoglobulins for many purposes, including increasing the valency of protein binding sites. For example, fragments of the proteins may be fused through "linker" sequences to the Fc portion of an immunoglobulin. For a bivalent form of the protein, such a fusion could be to the Fc portion of an IgG molecule. Other immunoglobulin

isotypes may also be used to generate such fusions. For example, a protein-IgM fusion would generate a decavalent form of the protein of the invention.

The present invention also provides both full-length and mature forms (for example, without a signal sequence or precursor sequence) of the disclosed proteins. The full-length form of the such proteins may be determined by translation of the nucleotide sequence of each disclosed clone. The mature form of such proteins may be obtained by expression of the disclosed full-length polynucleotide in a suitable mammalian cell or other host cell. The sequences of the mature form of the proteins are also determinable from the amino acid sequence of the full-length forms. Where proteins of the present invention are membrane bound, soluble forms of the proteins are also provided. In such forms part or all of the regions causing the proteins to be membrane bound are deleted so that the proteins are fully secreted from the cell in which it is expressed.

The present invention also provides genes corresponding to the cDNA sequences disclosed herein. The corresponding genes can be isolated in accordance with known methods using the sequence information disclosed herein. Such methods include the preparation of probes or primers from the disclosed sequence information for identification and/or amplification of genes in appropriate genomic libraries or other sources of genomic materials. Species homologs (e.g. orthologs) of the disclosed polynucleotides and proteins are also provided by the present invention. Species homologs may be isolated and identified by making suitable probes or primers from the sequences provided herein and screening a suitable nucleic acid source from the desired species. The invention also encompasses allelic variants of the disclosed polynucleotides or proteins; that is, naturally-occurring alternative forms of the isolated polynucleotide which also encode proteins which are identical, homologous or related to that encoded by the polynucleotides. The compositions of the present invention include isolated polynucleotides, including recombinant DNA molecules, cloned genes or degenerate variants thereof, especially naturally occurring variants such as allelic variants, novel isolated polypeptides, and antibodies that specifically recognize one or more epitopes present on such polypeptides. Species homologs of the disclosed polynucleotides and proteins are also provided by the present invention. Species homologs may be isolated and identified by making suitable probes or primers from the sequences provided herein and screening a suitable nucleic acid source from the desired species. The invention also

encompasses allelic variants of the disclosed polynucleotides or proteins; that is, naturally-occurring alternative forms of the isolated polynucleotide which also encode proteins which are identical, homologous or related to that encoded by the polynucleotides.

## 2. NUCLEIC ACIDS OF THE INVENTION

The isolated polynucleotides of the invention include, but are not limited to, a polynucleotide encoding a polypeptide comprising the amino acid sequence of SEQ ID NO: 2, 4, 6, 8, 11, 13, 15, 17, 19, 21, 23, 25, 27, 29, 31, 33, 35, 37, 39, 41, 43 or 45; or the mature protein portion thereof. A preferred nucleic acid sequence is set forth in SEQ ID NO: 1, 3, 5, 7, 9, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, 34, 36, 38, 40, 42 or 44 respectively.

The isolated polynucleotides of the invention further include, but are not limited to a polynucleotide comprising the nucleotide sequence of SEQ ID NO: 1, 3, 5, 7, 9, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, 34, 36, 38, 40, 42 or 44; a polynucleotide comprising the full length protein coding sequence of SEQ ID NO: 1, 3, 5, 7, 9, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, 34, 36, 38, 40, 42 or 44; and a polynucleotide comprising the nucleotide sequence of the mature protein coding sequence of SEQ ID NO: 1, 3, 5, 7, 9, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, 34, 36, 38, 40, 42 or 44. The polynucleotides of the present invention also include, but are not limited to, polynucleotides that encode polypeptides with biological activity and that hybridize under stringent hybridization conditions to the complement of either (a) the nucleotide sequence of SEQ ID NO:1, 3, 5, 7, 9, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, 34, 36, 38, 40, 42 or 44; or (b) a nucleotide sequence encoding the amino acid sequence of SEQ ID NO: 2, 4, 6, 8, 11, 13, 15, 17, 19, 21, 23, 25, 27, 29, 31, 33, 35, 37, 39, 41, 43 or 45; a polynucleotide which is an allelic variant of any polynucleotide recited above; a polynucleotide which encodes a species homolog of any of the proteins recited above; or a polynucleotide that encodes a polypeptide comprising a specific domain or truncation of the polypeptide of SEQ ID NO: 2, 4, 6, 8, 11, 13, 15, 17, 19, 21, 23, 25, 27, 29, 31, 33, 35, 37, 39, 41, 43 or 45.

The polynucleotides of the invention additionally include the complement of any of the polynucleotides described herein.

The polynucleotides of the invention also provide polynucleotides including nucleotide sequences that are substantially equivalent to the polynucleotides recited above. Polynucleotides according to the invention can have at least about 65%, more typically at least about 70%, at least about 75%, at least about 80%, at least about 85% or at least about 90%, and even more typically at least about 95%, sequence identity to a polynucleotide recited above. The invention also provides the complement of the polynucleotides including a nucleotide sequence that has at least about 80%, more typically at least about 90%, and even more typically at least about 95%, sequence identity to a polynucleotide encoding a polypeptide recited above. The polynucleotide can be DNA (genomic, cDNA, amplified, or synthetic) or RNA. Methods and algorithms for obtaining such polynucleotides are well known to those of skill in the art and can include, for example, methods for determining hybridization conditions which can routinely isolate polynucleotides of the desired sequence identities.

A polynucleotide according to the invention can be joined to any of a variety of other nucleotide sequences by well-established recombinant DNA techniques (see Sambrook J et al. (1989) Molecular Cloning: A Laboratory Manual, Cold Spring Harbor Laboratory, NY). Useful nucleotide sequences for joining to polypeptides include an assortment of vectors, e.g., plasmids, cosmids, lambda phage derivatives, phagemids, and the like, that are well known in the art. Accordingly, the invention also provides a vector including a polynucleotide of the invention and a host cell containing the polynucleotide. In general, the vector contains an origin of replication functional in at least one organism or host cell, convenient restriction endonuclease sites, and a selectable marker for the host cell. Vectors according to the invention include expression vectors, replication vectors, probe generation vectors, and sequencing vectors. A host cell according to the invention can be a prokaryotic or eukaryotic cell and can be a unicellular organism or part of a multicellular organism.

The sequences falling within the scope of the present invention are not limited to the specific sequences herein described, but also include allelic variations thereof. Allelic variations can be routinely determined by comparing the sequence provided in SEQ ID NO: 1, 3, 5, 7, 9, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, 34, 36, 38, 40, 42 or 44; or a representative fragment thereof; or a nucleotide sequence at least 99.9% identical to

SEQ ID NO: 1, 3, 5, 7, 9, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, 34, 36, 38, 40, 42 or 44, with a sequence from another isolate of the same species.

To accommodate codon variability, the invention includes nucleic acid molecules coding for the same amino acid sequences as do the specific ORFs disclosed herein. In other words, in the coding region of an ORF, substitution of one codon for another which encodes the same amino acid is expressly contemplated. Any specific sequence disclosed herein can be readily screened for errors by resequencing a particular fragment, such as an ORF, in both directions (i.e., sequence both strands).

The present invention further provides recombinant constructs comprising a nucleic acid having the sequence of SEQ ID NO: 1, 3, 5, 7, 9, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, 34, 36, 38, 40, 42 or 44; or a fragment thereof or any other polynucleotides of the invention. In one embodiment, the recombinant constructs of the present invention comprise a vector, such as a plasmid or viral vector, into which a nucleic acid having the sequence of SEQ ID NO: 1, 3, 5, 7, 9, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, 34, 36, 38, 40, 42 or 44; or a fragment thereof is inserted, in a forward or reverse orientation. In the case of a vector comprising one of the ORFs of the present invention, the vector may further comprise regulatory sequences, including for example, a promoter, operably linked to the ORF. For vectors comprising the EMFs and UMFs of the present invention, the vector may further comprise a marker sequence or heterologous ORF operably linked to the EMF or UMF. Large numbers of suitable vectors and promoters are known to those of skill in the art and are commercially available for generating the recombinant constructs of the present invention. The following vectors are provided by way of example. Bacterial: pBs, phagescript, PsiX174, pBluescript SK, pBs KS, pNH8a, pNH16a, pNH18a, pNH46a (Stratagene); pTrc99A, pKK223-3, pKK233-3, pDR540, pRIT5 (Pharmacia). Eukaryotic: pWLneo, pSV2cat, pOG44, PXTI, pSG (Stratagene) pSVK3, pBPV, pMSG, pSVL (Pharmacia).

The isolated polynucleotide of the invention may be operably linked to an expression control sequence such as the pMT2 or pED expression vectors disclosed in Kaufman et al., Nucleic Acids Res. 19, 4485-4490 (1991), in order to produce the protein recombinantly. Many suitable expression control sequences are known in the art. General methods of expressing recombinant proteins are also known and are exemplified in R. Kaufman, Methods in Enzymology 185, 537-566 (1990). As defined herein "operably

linked" means that the isolated polynucleotide of the invention and an expression control sequence are situated within a vector or cell in such a way that the protein is expressed by a host cell which has been transformed (transfected) with the ligated polynucleotide/expression control sequence.

Promoter regions can be selected from any desired gene using CAT (chloramphenicol transferase) vectors or other vectors with selectable markers. Two appropriate vectors are pKK232-8 and pCM7. Particular named bacterial promoters include lacI, lacZ, T3, T7, gpt, lambda $P_R$, and trc. Eukaryotic promoters include CMV immediate early, HSV thymidine kinase, early and late SV40 gene promoter, LTRs from retrovirus, and mouse metallothionein-I. Selection of the appropriate vector and promoter is well within the level of ordinary skill in the art. Generally, recombinant expression vectors will include origins of replication and selectable markers permitting transformation of the host cell, e.g., the ampicillin resistance gene of *E. coli* and *S. cerevisiae* TRP1 gene, and a promoter derived from a highly-expressed gene to direct transcription of a downstream structural sequence. Such promoters can be derived from operons encoding glycolytic enzymes such as 3-phosphoglycerate kinase (PGK), a-factor, acid phosphatase, or heat shock proteins, among others. The heterologous structural sequence is assembled in appropriate phase with translation initiation and termination sequences, and preferably, a leader sequence capable of directing secretion of translated protein into the periplasmic space or extracellular medium. Optionally, the heterologous sequence can encode a fusion protein including an N-terminal identification peptide imparting desired characteristics, e.g., stabilization or simplified purification of expressed recombinant product. Useful expression vectors for bacterial use are constructed by inserting a structural DNA sequence encoding a desired protein together with suitable translation initiation and termination signals in operable reading phase with a functional promoter. The vector will comprise one or more phenotypic selectable markers and an origin of replication to ensure maintenance of the vector and to, if desirable, provide amplification within the host. Suitable prokaryotic hosts for transformation include *E. coli, Bacillus subtilis, Salmonella typhimurium* and various species within the genera *Pseudomonas*, *Streptomyces*, and *Staphylococcus*, although others may also be employed as a matter of choice.

As a representative but non-limiting example, useful expression vectors for bacterial use can comprise a selectable marker and bacterial origin of replication derived from commercially available plasmids comprising genetic elements of the well known cloning vector pBR322 (ATCC 37017). Such commercial vectors include, for example, pKK223-3 (Pharmacia Fine Chemicals, Uppsala, Sweden) and GEM 1 (Promega Biotech, Madison, WI, USA). These pBR322 "backbone" sections are combined with an appropriate promoter and the structural sequence to be expressed. Following transformation of a suitable host strain and growth of the host strain to an appropriate cell density, the selected promoter is induced or derepressed by appropriate means (e.g., temperature shift or chemical induction) and cells are cultured for an additional period. Cells are typically harvested by centrifugation, disrupted by physical or chemical means, and the resulting crude extract retained for further purification.

Included within the scope of the nucleic acid sequences of the invention are nucleic acid sequences that hybridize under stringent conditions to a fragment of the DNA sequence of SEQ ID NO: 1, 3, 5, 7, 9, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, 34, 36, 38, 40, 42 or 44; which fragment is greater than about 10 bp, preferably 20-50 bp, greater than 100 bp, greater than 300 bp, or greater than 500 bp. In accordance with the invention, polynucleotide sequences which encode the novel nucleic acids, or functional equivalents thereof, may be used to generate recombinant DNA molecules that direct the expression of that nucleic acid, or a functional equivalent thereof, in appropriate host cells.

The nucleic acid sequences of the invention are further directed to sequences which encode variants of the described nucleic acids. These amino acid sequence variants may be prepared by methods known in the art by introducing appropriate nucleotide changes into a native or variant polynucleotide. There are two variables in the construction of amino acid sequence variants: the location of the mutation and the nature of the mutation. The amino acid sequence variants of the nucleic acids are preferably constructed by mutating the polynucleotide to give an amino acid sequence that does not occur in nature. These amino acid alterations can be made at sites that differ in the nucleic acids from different species or other family members (variable positions) or in highly conserved regions (constant regions). Sites at such locations will typically be modified in series, e.g., by substituting first with conservative choices (e.g., hydrophobic

amino acid to a different hydrophobic amino acid) and then with more distant choices (e.g., hydrophobic amino acid to a charged amino acid), and then deletions or insertions may be made at the target site. Amino acid sequence deletions generally range from about 1 to 30 residues, preferably about 1 to 10 residues, and are typically contiguous. Amino acid insertions include amino- and/or carboxyl-terminal fusions ranging in length from one to one hundred or more residues, as well as intrasequence insertions of single or multiple amino acid residues. Intrasequence insertions may range generally from about 1 to 10 amino residues, preferably from 1 to 5 residues. Examples of terminal insertions include the heterologous signal sequences necessary for secretion or for intracellular targeting in different host cells, and sequences such as FLAG or poly-histidine sequences useful for purifying the expressed protein..

In a preferred method, polynucleotides encoding the novel nucleic acids are changed via site-directed mutagenesis. This method uses oligonucleotide sequences that encode the polynucleotide sequence of the desired amino acid variant, as well as a sufficient adjacent nucleotide on both sides of the changed amino acid to form a stable duplex on either side of the site of being changed. In general, the techniques of site-directed mutagenesis are well known to those of skill in the art and this technique is exemplified by publications such as, Edelman et al., DNA 2:183 (1983). A versatile and efficient method for producing site-specific changes in a polynucleotide sequence was published by Zoller and Smith, Nucleic Acids Res. 10:6487-6500 (1982). PCR may also be used to create amino acid sequence variants of the novel nucleic acids. When small amounts of template DNA are used as starting material, primer(s) that differs slightly in sequence from the corresponding region in the template DNA can generate the desired amino acid variant. PCR amplification results in a population of product DNA fragments that differ from the polynucleotide template encoding the polypeptide at the position specified by the primer. The product DNA fragments replace the corresponding region in the plasmid and this gives the desired amino acid variant.

A further technique for generating amino acid variants is the cassette mutagenesis technique described in Wells et al., Gene 34:315 (1985); and other mutagenesis techniques well known in the art, such as, for example, the techniques in Sambrook et al., supra, and Current Protocols in Molecular Biology, Ausubel et al. Due to the inherent degeneracy of the genetic code, other DNA sequences which encode substantially the

same or a functionally equivalent amino acid sequence may be used in the practice of the invention for the cloning and expression of these novel nucleic acids. Such DNA sequences include those which are capable of hybridizing to the appropriate novel nucleic acid sequence under stringent conditions.

**3. HOSTS**

The present invention further provides host cells genetically engineered to contain the polynucleotides of the invention. For example, such host cells may contain nucleic acids of the invention introduced into the host cell using known transformation, transfection or infection methods. The present invention still further provides host cells genetically engineered to express the polynucleotides of the invention, wherein such polynucleotides are in operative association with a regulatory sequence heterologous to the host cell which drives expression of the polynucleotides in the cell.

Knowledge of DNA sequences provided by the invention (e.g. DNA encoding apolipoprotein, lipase, or lipoprotein receptor polypeptides of the invention) allows for modification of cells to permit, or increase, expression of endogenous polypeptide. Cells can be modified (*e.g.*, by homologous recombination) to provide increased polypeptide expression by replacing, in whole or in part, the naturally occurring promoter with all or part of a heterologous promoter so that the cells express the protein at higher levels. The heterologous promoter is inserted in such a manner that it is operatively linked to the desired protein encoding sequences. See, for example, PCT International Publication No. WO 94/12650, PCT International Publication No. WO 92/20808, and PCT International Publication No. WO 91/09955. It is also contemplated that, in addition to heterologous promoter DNA, amplifiable marker DNA (*e.g.*, ada, dhfr, and the multifunctional CAD gene which encodes carbamyl phosphate synthase, aspartate transcarbamylase, and dihydroorotase) and/or intron DNA may be inserted along with the heterologous promoter DNA. If linked to the desired protein coding sequence, amplification of the marker DNA by standard selection methods results in co-amplification of the desired protein coding sequences in the cells.

The host cell can be a higher eukaryotic host cell, such as a mammalian cell, a lower eukaryotic host cell, such as a yeast cell, or the host cell can be a prokaryotic cell, such as a bacterial cell. Introduction of the recombinant construct into the host cell can be

effected by calcium phosphate transfection, DEAE dextran mediated transfection, or electroporation (Davis, L. et al., Basic Methods in Molecular Biology (1986)). The host cells containing one of the polynucleotides of the invention, can be used in conventional manners to produce the gene product encoded by the isolated fragment (in the case of an ORF) or can be used to produce a heterologous protein under the control of the EMF.

Any host/vector system can be used to express one or more of the ORFs of the present invention. These include, but are not limited to, eukaryotic hosts such as HeLa cells, Cv-1 cell, COS cells, and Sf9 cells, as well as prokaryotic host such as *E. coli* and *B. subtilis*. The most preferred cells are those which do not normally express the particular polypeptide or protein or which expresses the polypeptide or protein at low natural level. Mature proteins can be expressed in mammalian cells, yeast, bacteria, or other cells under the control of appropriate promoters. Cell-free translation systems can also be employed to produce such proteins using RNAs derived from the DNA constructs of the present invention. Appropriate cloning and expression vectors for use with prokaryotic and eukaryotic hosts are described by Sambrook, et al., in Molecular Cloning: A Laboratory Manual, Second Edition, Cold Spring Harbor, New York (1989), the disclosure of which is hereby incorporated by reference.

Various mammalian cell culture systems can also be employed to express recombinant protein. Examples of mammalian expression systems include the COS-7 lines of monkey kidney fibroblasts, described by Gluzman, Cell 23:175 (1981), and other cell lines capable of expressing a compatible vector, for example, the C127, 3T3, CHO, HeLa and BHK cell lines. Mammalian expression vectors will comprise an origin of replication, a suitable promoter and also any necessary ribosome binding sites, polyadenylation site, splice donor and acceptor sites, transcriptional termination sequences, and 5' flanking nontranscribed sequences. DNA sequences derived from the SV40 viral genome, for example, SV40 origin, early promoter, enhancer, splice, and polyadenylation sites may be used to provide the required nontranscribed genetic elements. Recombinant polypeptides and proteins produced in bacterial culture are usually isolated by initial extraction from cell pellets, followed by one or more salting-out, aqueous ion exchange or size exclusion chromatography steps. Protein refolding steps can be used, as necessary, in completing configuration of the mature protein. Finally, high performance liquid chromatography (HPLC) can be employed for

final purification steps. Microbial cells employed in expression of proteins can be disrupted by any convenient method, including freeze-thaw cycling, sonication, mechanical disruption, or use of cell lysing agents.

A number of types of cells may act as suitable host cells for expression of the protein. Mammalian host cells include, for example, monkey COS cells, Chinese Hamster Ovary (CHO) cells, human kidney 293 cells, human epidermal A431 cells, human Colo205 cells, 3T3 cells, CV-1 cells, other transformed primate cell lines, normal diploid cells, cell strains derived from in vitro culture of primary tissue, primary explants, HeLa cells, mouse L cells, BHK, HL-60, U937, HaK or Jurkat cells.

Alternatively, it may be possible to produce the protein in lower eukaryotes such as yeast, insects or in prokaryotes such as bacteria. Potentially suitable yeast strains include *Saccharomyces cerevisiae*, *Schizosaccharomyces pombe*, *Kluyveromyces* strains, *Candida*, or any yeast strain capable of expressing heterologous proteins. Potentially suitable bacterial strains include *Escherichia coli*, *Bacillus subtilis*, *Salmonella typhimurium*, or any bacterial strain capable of expressing heterologous proteins. If the protein is made in yeast or bacteria, it may be necessary to modify the protein produced therein, for example by phosphorylation or glycosylation of the appropriate sites, in order to obtain the functional protein. Such covalent attachments may be accomplished using known chemical or enzymatic methods.

In another embodiment of the present invention, cells and tissues may be engineered to express an endogenous gene comprising the polynucleotides of the invention under the control of inducible regulatory elements, in which case the regulatory sequences of the endogenous gene may be replaced by homologous recombination. As described herein, gene targeting can be used to replace a gene's existing regulatory region with a regulatory sequence isolated from a different gene or a novel regulatory sequence synthesized by genetic engineering methods. Such regulatory sequences may be comprised of promoters, enhancers, scaffold-attachment regions, negative regulatory elements, transcriptional initiation sites, regulatory protein binding sites or combinations of said sequences. Alternatively, sequences which affect the structure or stability of the RNA or protein produced may be replaced, removed, added, or otherwise modified by targeting, including polyadenylation signals. mRNA stability elements, splice sites, leader sequences for enhancing or modifying transport or secretion properties of the

protein, or other sequences which alter or improve the function or stability of protein or RNA molecules.

The targeting event may be a simple insertion of the regulatory sequence, placing the gene under the control of the new regulatory sequence, e.g., inserting a new promoter or enhancer or both upstream of a gene. Alternatively, the targeting event may be a simple deletion of a regulatory element, such as the deletion of a tissue-specific negative regulatory element. Alternatively, the targeting event may replace an existing element; for example, a tissue-specific enhancer can be replaced by an enhancer that has broader or different cell-type specificity than the naturally occurring elements. Here, the naturally occurring sequences are deleted and new sequences are added. In all cases, the identification of the targeting event may be facilitated by the use of one or more selectable marker genes that are contiguous with the targeting DNA, allowing for the selection of cells in which the exogenous DNA has integrated into the host cell genome. The identification of the targeting event may also be facilitated by the use of one or more marker genes exhibiting the property of negative selection, such that the negatively selectable marker is linked to the exogenous DNA, but configured such that the negatively selectable marker flanks the targeting sequence, and such that a correct homologous recombination event with sequences in the host cell genome does not result in the stable integration of the negatively selectable marker. Markers useful for this purpose include the Herpes Simplex Virus thymidine kinase (TK) gene or the bacterial xanthine-guanine phosphoribosyl-transferase (gpt) gene.

Exemplary gene targeting or gene activation techniques which can be used in accordance with this aspect of the invention are more particularly described in U.S. Patent No. 5,272,071 to Chappel; U.S. Patent No. 5,578,461 to Sherwin et al.; International Application No. PCT/US92/09627 (WO93/09222) by Selden et al.; and International Application No. PCT/US90/06436 (WO91/06667) by Skoultchi et al., each of which is incorporated by reference herein in its entirety.

**4. POLYPEPTIDES OF THE INVENTION**

The isolated polypeptides of the invention include, but are not limited to, a polypeptide comprising the amino acid sequence of SEQ ID NO: 2, 4, 6, 8, 11, 13, 15, 17, 19, 21, 23, 25, 27, 29, 31, 33, 35, 37, 39, 41, 43 or 45; or the amino acid sequence

encoded by the DNA of SEQ ID NO: 1, 3, 5, 7, 9, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, 34, 36, 38, 40, 42 or 44; or the corresponding full length or mature protein. Polypeptides of the invention also include polypeptides preferably with biological or immunological activity that are encoded by (a) the polynucleotide of SEQ ID NO: 1, 3, 5, 7, 9, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, 34, 36, 38, 40, 42 or 44; or (b) polynucleotides encoding SEQ ID NO: 2, 4, 6, 8, 11, 13, 15, 17, 19, 21, 23, 25, 27, 29, 31, 33, 35. 37, 39, 41, 43 or 45; or polynucleotides that hybridize to the complement of the polynucleotides of either (a) or (b) under stringent hybridization conditions. Biologically active or immunologically active variants of the polypeptide amino acid sequence of SEQ ID NO: 2, 4, 6, 8, 11, 13, 15, 17, 19, 21, 23, 25, 27, 29, 31, 33, 35, 37, 39, 41, 43 or 45, or the corresponding full length or mature protein; and "substantial equivalents" thereof (*e.g.*, with 65%, 70%, 75%, 80%, 85%, 90%, typically 95%, more typically 98%, or most typically 99% amino acid identity) that retain a biological activity, preferably apoprotein, lipase, or lipoprotein receptor activity are contemplated. Polypeptides encoded by allelic variants may have a similar, increased, or decreased activity compared to polypeptides of SEQ ID NOS: 2, 4, 6, 8, 11, 13, 15, 17, 19, 21, 23, 25, 27, 29, 31, 33, 35, 37, 39, 41, 43 or 45.

Protein compositions of the present invention may further comprise an acceptable carrier, such as a hydrophilic, e.g., pharmaceutically acceptable, carrier.

The invention also relates to methods for producing a polypeptide comprising growing a culture of the cells of the invention in a suitable culture medium, and purifying the protein from the cells or the culture in which the cells are grown. For example, the methods of the invention include a process for producing a polypeptide in which a host cell containing a suitable expression vector that includes a polynucleotide of the invention is cultured under conditions that allow expression of the encoded polypeptide. The polypeptide can be recovered from the cells or the culture medium, and further purified. Preferred embodiments include those in which the protein produced by such process is a full length or mature form of the protein.

The present invention further provides isolated polypeptides encoded by the nucleic acid fragments of the present invention or by degenerate variants of the nucleic acid fragments of the present invention. By "degenerate variant" is intended nucleotide fragments which differ from a nucleic acid fragment of the present invention (e.g., an

ORF) by nucleotide sequence due to the degeneracy of the genetic code, but which encode an identical polypeptide sequence. Preferred nucleic acid fragments of the present invention are the ORFs that encode proteins. A variety of methodologies known in the art can be utilized to obtain any one of the isolated polypeptides or proteins of the present invention. At the simplest level, the amino acid sequence can be synthesized using commercially available peptide synthesizers. This is particularly useful in producing small peptides and fragments of larger polypeptides. Fragments are useful, for example, in generating antibodies against the native polypeptide. In an alternative method, the polypeptide or protein is purified from host cells which produce the polypeptide or protein. One skilled in the art can readily follow known methods for isolating polypeptides and proteins in order to obtain one of the isolated polypeptides or proteins of the present invention. These include, but are not limited to, immunochromatography, HPLC, size-exclusion chromatography, ion-exchange chromatography, and immuno-affinity chromatography. See, e.g., Scopes, Protein Purification: Principles and Practice, Springer-Verlag (1994); Sambrook, et al., in Molecular Cloning: A Laboratory Manual; Ausubel et al., Current Protocols in Molecular Biology. Polypeptide fragments that retain biological/immunological activity include fragments encoding greater than about 100 amino acids, or greater than about 200 amino acids, and fragments that encode specific protein domains.

The polypeptides and proteins of the present invention can alternatively be purified from cells which have been altered to express the desired polypeptide or protein. As used herein, a cell is said to be altered to express a desired polypeptide or protein when the cell, through genetic manipulation, is made to produce a polypeptide or protein which it normally does not produce or which the cell normally produces at a lower level. One skilled in the art can readily adapt procedures for introducing and expressing either recombinant or synthetic sequences into eukaryotic or prokaryotic cells in order to generate a cell which produces one of the polypeptides or proteins of the present invention. The purified polypeptides can be used in *in vitro* binding assays which are well known in the art to identify molecules which bind to the polypeptides.

Sources for test compounds that may be screened for ability to bind to or modulate (i.e., increase or decrease) the activity of polypeptides of the invention include (1) inorganic and organic chemical libraries, (2) natural product libraries, and (3)

combinatorial libraries comprised of either random or mimetic peptides, oligonucleotides or organic molecules.

Chemical libraries may be readily synthesized or purchased from a number of commercial sources, and may include structural analogs of known compounds or compounds that are identified as "hits" or "leads" via natural product screening.

The sources of natural product libraries are microorganisms (including bacteria and fungi), animals, plants or other vegetation, or marine organisms, and libraries of mixtures for screening may be created by: (1) fermentation and extraction of broths from soil, plant or marine microorganisms or (2) extraction of the organisms themselves. Natural product libraries include polyketides, non-ribosomal peptides, and (non-naturally occurring) variants thereof. For a review, see *Science 282*:63-68 (1998).

Combinatorial libraries are composed of large numbers of peptides, oligonucleotides or organic compounds and can be readily prepared by traditional automated synthesis methods, PCR, cloning or proprietary synthetic methods. Of particular interest are peptide and oligonucleotide combinatorial libraries. Still other libraries of interest include peptide, protein, peptidomimetic, multiparallel synthetic collection, recombinatorial, and polypeptide libraries. For a review of combinatorial chemistry and libraries created therefrom, see Myers, *Curr. Opin. Biotechnol.* 8:701-707 (1997). For reviews and examples of peptidomimetic libraries, see Al-Obeidi et al., *Mol. Biotechnol*, 9(3):205-23 (1998); Hruby et al., *Curr Opin Chem Biol*, 1(1):114-19 (1997); Dorner et al., *Bioorg Med Chem*, 4(5):709-15 (1996) (alkylated dipeptides).

Identification of modulators through use of the various libraries described herein permits modification of the candidate "hit" (or "lead") to optimize the capacity of the "hit" to bind a polypeptide of the invention. The molecules identified in the binding assay are then tested for antagonist or agonist activity in *in vivo* tissue culture or animal models that are well known in the art. In brief, the molecules are titrated into a plurality of cell cultures or animals and then tested for either cell/animal death or prolonged survival of the animal/cells.

In addition, the binding molecules may be complexed with toxins, e.g., ricin or cholera, or with other compounds that are toxic to cells such as radioisotopes. The toxin-binding molecule complex is then targeted to a tumor or other cell by the specificity of the binding molecule for a polypeptide of the invention. Alternatively, the polypeptide

of the invention or binding molecules may be complexed with imaging agents for targeting and imaging purposes.

The protein of the invention may also be expressed as a product of transgenic animals, e.g., as a component of the milk of transgenic cows, goats, pigs, or sheep which are characterized by somatic or germ cells containing a nucleotide sequence encoding the protein.

The protein may also be produced by known conventional chemical synthesis. Methods for constructing the proteins of the present invention by synthetic means are known to those skilled in the art. The synthetically-constructed protein sequences, by virtue of sharing primary, secondary or tertiary structural and/or conformational characteristics with proteins may possess biological properties in common therewith, including protein activity. Thus, they may be employed as biologically active or immunological substitutes for natural, purified proteins in screening of therapeutic compounds and in immunological processes for the development of antibodies.

The proteins provided herein also include proteins characterized by amino acid sequences similar to those of purified proteins but into which modification are naturally provided or deliberately engineered. For example, modifications in the peptide or DNA sequences can be made by those skilled in the art using known techniques. Modifications of interest in the protein sequences may include the alteration, substitution, replacement, insertion or deletion of a selected amino acid residue in the coding sequence. For example, one or more of the cysteine residues may be deleted or replaced with another amino acid to alter the conformation of the molecule. Techniques for such alteration, substitution, replacement, insertion or deletion are well known to those skilled in the art (see, e.g., U.S. Pat. No. 4,518,584). Preferably, such alteration, substitution, replacement, insertion or deletion retains a desired activity of the protein.

Other fragments and derivatives of the sequences of proteins which would be expected to retain protein activity in whole or in part and may thus be useful for screening or other immunological methodologies may also be easily made by those skilled in the art given the disclosures herein. Such modifications are believed to be encompassed by the present invention.

The protein may also be produced by operably linking the isolated polynucleotide of the invention to suitable control sequences in one or more insect expression vectors,

and employing an insect expression system. Materials and methods for baculovirus/insect cell expression systems are commercially available in kit form from, e.g., Invitrogen, San Diego, Calif., U.S.A. (the MaxBat.RTM. kit), and such methods are well known in the art, as described in Summers and Smith, Texas Agricultural Experiment Station Bulletin No. 1555 (1987), incorporated herein by reference. As used herein, an insect cell capable of expressing a polynucleotide of the present invention is "transformed."

The protein of the invention may be prepared by culturing transformed host cells under culture conditions suitable to express the recombinant protein. The resulting expressed protein may then be purified from such culture (i.e., from culture medium or cell extracts) using known purification processes, such as gel filtration and ion exchange chromatography. The purification of the protein may also include an affinity column containing agents which will bind to the protein; one or more column steps over such affinity resins as concanavalin A-agarose, heparin-toyopearl™ or Cibacrom blue 3GA Sepharose™; one or more steps involving hydrophobic interaction chromatography using such resins as phenyl ether, butyl ether, or propyl ether; or immunoaffinity chromatography.

Alternatively, the protein of the invention may also be expressed in a form which will facilitate purification. For example, it may be expressed as a fusion protein, such as fused with maltose binding protein (MBP), glutathione-S-transferase (GST), thioredoxin (TRX), or as a His tag. Kits for expression and purification of such fusion proteins are commercially available from New England BioLab (Beverly, Mass.), Pharmacia (Piscataway, N.J.), Invitrogen, and Qiagen respectively. The protein can also be tagged with an epitope and subsequently purified by using a specific antibody directed to such epitope. One such epitope ("Flag") is commercially available from Kodak (New Haven, Conn.).

Finally, one or more reverse-phase high performance liquid chromatography (RP-HPLC) steps employing hydrophobic RP-HPLC media, e.g., silica gel having pendant methyl or other aliphatic groups, can be employed to further purify the protein. Some or all of the foregoing purification steps, in various combinations, can also be employed to provide a substantially homogeneous isolated recombinant protein. The protein thus purified is substantially free of other mammalian proteins and is defined in accordance with the present invention as an "isolated protein."

The polypeptides of the invention include CG122, CG179, CG95, CG121, CG162, CG27, CG153, and CG168 analogs (variants). This embraces fragments of CG122, CG179, CG95, CG121, CG162, CG27, CG153, and CG168; as well as analogs (variants) thereof in which one or more amino acids has been deleted, inserted, or substituted. Analogs of the invention also embrace fusions or modifications of CG122, CG179, CG95, CG121, CG162, CG27, CG153, and CG168; wherein the protein or analog is fused to another moiety or moieties, e.g., targeting moiety or another therapeutic agent. Such analogs may exhibit improved properties such as activity and/or stability. Examples of moieties which may be fused to CG122, CG179, CG95, CG121, CG162, CG27, CG153, CG168 or an analog include, for example, targeting moieties which provide for the delivery of polypeptide to desired cell types. Other moieties which may be fused to CG122, CG179, CG95, CG121, CG162, CG27, CG153, CG168 or an analog include therapeutic agents which are used for treatment of disorders described herein.

**5. GENE THERAPY**

Mutations in the CG122, CG179, CG95, CG121, CG162, CG27, CG153 or CG168 gene may result in loss of normal function of the encoded protein. The invention thus provides gene therapy to restore normal CG122, CG179, CG95, CG121, CG162, CG27, CG153 or CG168 activity; or to treat disease states involving CG122, CG179, CG95, CG121, CG162, CG27, CG153 or CG168. Delivery of a functional CG122, CG179, CG95, CG121, CG162, CG27, CG153 or CG168 gene to appropriate cells is effected *ex vivo*, *in situ*, or *in vivo* by use of vectors, and more particularly viral vectors (*e.g.*, adenovirus, adeno-associated virus, or a retrovirus), or *ex vivo* by use of physical DNA transfer methods (e.g., liposomes or chemical treatments). See, for example, Anderson, Nature, supplement to vol. 392, no. 6679, pp.25-20 (1998). For additional reviews of gene therapy technology see Friedmann, Science, 244: 1275-1281 (1989); Verma, Scientific American: 68-84 (1990); and Miller, Nature, 357: 455-460 (1992). Alternatively, it is contemplated that in other human disease states, preventing the expression of or inhibiting the activity of CG122, CG179, CG95, CG121, CG162, CG27, CG153 or CG168 will be useful in treating the disease states. It is contemplated that antisense therapy or gene therapy could be applied to negatively regulate the expression of CG122, CG179, CG95, CG121, CG162, CG27, CG153 or CG168.

**5.1 TRANSGENIC ANIMALS**

With a polynucleotide of the invention, transgenic animals can be produced wherein a polynucleotide encoding the desired specific binding agent is introduced into the genome of a recipient animal in a manner that permits expression of the encoded specific binding agent, or alternatively, the sequence in an animal can be disabled so that at least one allele in nonfunctional. Two methods of producing transgenic mice are widely used. In one method, embryonic stem cells (ES cells) in tissue culture are transformed with a desired DNA, and in an alternative method, a desired polynucleotide is injected into the pronucleus of a fertilized mouse egg.

In the first method, ES cells are harvested from the inner cell mass of mouse blastocysts. The isolated cells can be grown in culture and generally retain their full potential to produce all the cells of the mature animal. Cells growing in culture are transformed/transfected by methods well known and routinely used in the art, and cells are selected based generally on expression of some marker encoded by the transforming DNA (see below). Selected cells are then injected into the inner cell mass (ICM) of mouse blastocyst. These embryos are transferred to the uterus of a pseudo pregnant mouse (produced by mating a female mouse with a vasectomized male). Offspring are then tested by removing a small piece of tissue from the tail and examine its DNA for the desired gene and offspring that are found to have the desired DNA will be heterozygous. A homozygous strain can then be produced by mating two heterozygotes.

In the second method freshly fertilized eggs are harvested before the sperm head becomes a pronucleus. Desired DNA is injected into the male pronucleus and when the pronuclei have fused to form the diploid zygote nucleus, the zygote is allowed to form a 2-cell embryo. These embryos are then implanted in a pseudopregnant mouse as described above and resulting offspring examined, also as described.

The design of the DNA used in these methods is based on the desired results, including, for example, restoring gene function in a mutant animal or knocking out the function of a particular locus. In either case, the designed DNA will include the targeted gene insertion, and generally *neo*$^r$, a selectable marker gene that encodes an enzyme that inactivates the antibiotic neomycin (and its relatives) and/or *tk*, a gene that encodes thymidine kinase, an enzyme that phosphorylates the nucleoside analog gancyclovir. DNA polymerase fails to discriminate against the resulting nucleotide and inserts this

nonfunctional nucleotide into freshly-replicating DNA which is generally lethal to the cell. Following random insertion, the entire vector, including the *tk* gene, is stably integrated into the host genome and the resulting cells are resistant to neomycin but killed by gancyclovir. In some cells, homologous recombination will occur wherein only part of the designed DNA will stably insert into the host genome. Cells are therefore first selected by culturing the cells in neomycin; cells that failed to take up the vector are killed. A second selection includes culturing the selected cells in gancyclovir which will identify those cells transformed by homologous recombination. These cells are then injected into the inner cell mass of mouse blastocyst as described above. Other selectable markers are well known in the art and can be utilized in place of those described herein. these methods.

When the transforming DNA is nonfunctional (for example, in the production of knockout animals to produce a "null" allele), the resulting offspring will be heterozygous. Mating of heterozygous transgenic animals, however, will produce a strain of "knockouts" homozygous for the null allele gene. In general, transgenic animals are produced using mice.

Alternatively, sheep fibroblasts growing can be grown in tissue culture and transformed or transfected DNA as described above, including, for example, a neomycin-resistance gene to aid in selection, and a desired gene sequence under control of one or more promoter sites from the beta-lactoglobulin gene. Integration of this chimeric gene permits expression in milk-producing cells. Successfully-transformed cells can be fused with enucleated sheep eggs and implanted in the uterus of a ewe. Surviving offspring are expected to produce the desired protein in milk. See, Pollock, *et al.*, J. Immunol. Meth. 231:147-157 (1999); Little, *et al.*, *Immunol. Today* 8: 364-370 (2000). The protein of the invention may also be expressed as a product of transgenic animals, and particularly as a component of the milk of transgenic cows, goats, or pigs, which are characterized by somatic or germ cells containing a nucleotide sequence encoding the protein.

In methods to determine biological functions of CG122, CG179, CG95, CG121, CG162, CG27, CG153, and CG168, *in vivo*, one or more genes provided by the invention are either over expressed or inactivated in the germ line of animals using homologous recombination [Capecchi, Science 244:1288-1292 (1989)]. Animals in which the gene is

over expressed, under the regulatory control of exogenous or endogenous promoter elements, are known as transgenic animals. Animals in which an endogenous gene has been inactivated by homologous recombination are referred to as "knockout" animals. Knockout animals, preferably non-human mammals, can be prepared as described in U.S. Patent No. 5,557,032, incorporated herein by reference. Such transgenic animals are useful to determine the roles CG122, CG179, CG95, CG121, CG162, CG27, CG153, and CG168 play in biological processes, and preferably in disease states. Transgenic animals are useful as model systems to identify compounds that modulate lipid metabolism. Transgenic animals, preferably non-human mammals, are produced using methods as described in U.S. Patent No 5,489,743 and PCT Publication No. WO94/28122, incorporated herein by reference.

Transgenic animals can be prepared wherein all or part of an CG122, CG179, CG95, CG121, CG162, CG27, CG153 or CG168 promoter is either activated or inactivated to alter the level of expression of the CG122, CG179, CG95, CG121, CG162, CG27, CG153 or CG168 protein. Inactivation can be carried out using homologous recombination methods described above. Activation can be achieved by supplementing or even replacing the homologous promoter to provide for increased protein expression. The homologous promoter can be supplemented by insertion of one or more heterologous enhancer elements known to confer promoter activity in a particular tissue. The promoter may also be introduced into functional proximity to the recited genes by homologous recombination.

## 6. USES AND BIOLOGICAL ACTIVITY

The biological activity of a polypeptide of the invention may manifest as, *e.g.*, apolipoprotein, lipase, or lipoprotein receptor signaling activity. The polynucleotides and proteins of the present invention are expected to exhibit one or more of the uses or biological activities (including those associated with assays cited herein) identified below. Uses or activities described for proteins of the present invention may be provided by administration or use of such proteins or by administration or use of polynucleotides encoding such proteins (such as, for example, in gene therapies or vectors suitable for introduction of DNA). The mechanism underlying the particular condition or pathology will dictate whether CG122, CG179, CG95, CG121, CG162, CG27, CG153 or CG168

polypeptides; polynucleotides; or modulators (activators and inhibitors) would be beneficial to the subject in need of treatment. Thus, "therapeutic compositions of the invention" include compositions comprising of polynucleotides or polypeptides of the invention or compounds and other substances that modulate the overall activity of the target CG122, CG179, CG95, CG121, CG162, CG27, CG153 or CG168 gene products, either at the level of target gene/protein expression or target protein activity. Such modulators include polypeptides, analogs, (variants), including fragments and fusion proteins, antibodies and other binding proteins; compounds that directly or indirectly activate or inhibit the apolipoprotein-like, lipase-like, or lipoprotein receptor-like polypeptides of the invention; and antisense polynucleotides and polynucleotides suitable for triple helix formation.

CG122 and CG179 are related to members of the apolipoprotein family which include apo AI, A-II, A-IV, B, CI, CII, CIII, D, E, H, J, L, and apo(a), among others. CG122 most closely resembles apo IV while CG179 is most similar to apo C. CG95, CG121, and CG162 are all putative lipases. CG95 shown greatest similarity to $PLA_2$, CG121 to PLC, and CG162 to LAL. CG27, CG153, and CG168 are related to the lipoprotein receptors LDL receptor, VLDL receptor, scavenger receptor, and LRP respectively.

Changes in lipoprotein metabolism that lead to atherosclerosis or coronary heart disease can be due to diet or to mutations in genes encoding proteins involved in lipid transport [Breslow (1993) Circ 87 suppl III: III-16-III-21]. A number of such mutations are found in genes encoding the apolipoprotein component of lipoproteins. For example, abnormalities in apo E lead to type III hyperlipidemias also known as dysbetalipoproteinemia. Mutations in apo B can cause heterozygous hypobetalipoproteinemia or familial defective apo B-100. Defects in apo A-I can lead to very low HDL cholesterol levels and premature coronary heart disease, or to the apo A-$I_{Milano}$ disorder [Breslow (1993) Circ 87 suppl III: III-16-III-21; Beiseigel (1998) Eur Heart J Suppl A: A20-A23]. Defects in other proteins that regulate lipid metabolisn such as LPL can lead to massive hyperglyceridaemias such as chylomicronaemias, mixed hyperlipidaemia, postprandial hyperlipidaemias, and to low HDL. Mutations in the LDL receptor can lead to severe hypercholesterolaemia. Tangier disease, caused by mutations in ABC1 (also known as CERP) causes abnormalities in cholesterol metabolism and can

lead to premature coronary artery disease [Rust et al. (1999) Nat Genet 22:352-355; Brooks-Wilson et al. (1999) Nat Genet 22:336-345]]. Defects in LAL activity, important for the regulation of cellular lipid uptake, is the underlying cause of two heritable diseases: Wolman disease and cholesteryl ester storage disease (CESD). Some pateints with CESD are able to survive past middle age but show signs of premature atherosclerosis [Du et al. (1998) Mol Gen Meta 64:126-134]. Other disorders, such as hypertriglyceridemia, may also result from defects in proteins involved in lipid metabolism [Breslow (1993) Circ 87 suppl III: III-16-III-21; Beiseigel (1998) Eur Heart J Suppl A: A20-A23].

Increased levels of extracellular snp$PLA_2$ activity has been associated with numerous inflammatory conditions including atherosclerosis and other cardiovascular diseases. snp$PLA_2$ is found associated with SMCs in normal arteries as well as the intima of atherosclerotic arteries, macrophages, and the lipid core of atherosclerotic plaques. snp$PLA_2$ is anchored to the extracellular matrix of arterial walls by binding to sulfated glycosaminoglycans (GAG) on proteoglycans. Chondroitin-sulfate proteoglycans (CSPG), such as versican, is expressed in the tunica of normal arteries and in the intima of atherosclerotic arteries. LDL and snp$PLA_2$ are both bound to CSPGs bringing these molecules close together thus facilitating the rapid hydrolysis of LDL phospholipids into the pro-inflammatory lipid factors, FFA and lysophospholipids. This process decreases the number of phospholipids on the surface of LDL. Smaller LDL particles show greater affinity for GAG which prolongs the retention time of these lipoproteins in the arterial wall, thereby promoting and sustaining inflammatory responses in atherosclerotic lesions [Hurt-Camejo et al. (1997) Atherosclerosis 132:1-8].

The cytosolic phospholipase C family of enzymes include ten different mammalian isozymes that comprise three major subfamilies, PLC-$\beta$, PLC-$\gamma$, and PLC-$\delta$. PLC-$\gamma$ differs from the other members by inclusion of SH domains that mediate protein-protein interactions. PLC-$\gamma$ is an intracellular signaling molecule which is stimulated by a variety of agonists including *e.g.* hormones, growth factors, etc., that mediates the hydrolysis of phophatidylinositol 4,5-bisphosphate ($PIP_2$) into the second messengers, inositol 1,4,5-trisphosphate ($IP_3$) and 1,2-diacylglycerol (DAG). $IP_3$ induces the release of intracellular Ca2+ ions and DAG activates protein kinase C (PKC) leading to number of

different downstream cellular responses [Sekiya et a. (1999) Chem Phy Lip 98:3-11]. $PIP_2$ is also one of the activators of cytosolic phopholipase $A_2$ ($cPLA_2$). $cPLA_2$ is a member of a group of $PLA_2$ enzymes which also include calcium-independent $PLA_2$ ($iPLA_2$), and several secreted $PLA_2$s ($sPLA_2$). $cPLA_2$ releases arachidonic acid from membrane phospholipids such as 1-alkyl-2-archidonoyl-*sn*-glycero-3-phosphocholine, into the cytoplasm, in response to various stimuli that increase intracellular $Ca^{2+}$ ion concentration and lead to the phosphorylation of $cPLA_2$ via the MAP kinase pathway. Arachidonic acid is the precursor of pro-inflammatory lipids which include the eicosanoids: leukotrienes, prostaglandins, and thromboxanes. Analysis of $cPLA_2$-deficient mice reveals that loss of this protein leads to a significant decrease in eicosanoid production revealing the important role of this protein in inflammatory responses. [Gijon et al. (1999) J Leuk Biol 65:330-336; Bayon et a. (1998) Cyto Cell Mol Therapy 4:275-286; Chaminade et al. (1999) Lipids 34 Suppl.:S49-S55].

Receptors that may be involved in the process of lipid accumulation include scavenger receptors expressed on macrophages and endothelial cells, and LRP and VLDL receptors expressed on SMCs [Greaves et al. (1998) Curr Opin Lipidol 9:425-432; Ylä-Herttuala (1996) Curr Opin Lipidol 7:292-297; Freeman (1997) Curr Opin Hematology 4:41-47]. Recent identification of scavenger receptors expressed by endothelial cells suggests that this cell type may also be involved in atherogenesis [Greaves et al. (1998) Curr Opin Lipidol 9:425-432; Hiltunen et al. (1998) Atherosclerosis 137 Suppl:S81-S88].

The LDL receptor gene family includes LDL receptor, VLDL receptor, LRP, LRP-2/Gp330/megalin, apoER2 or LR7/8B, and LR11/sorLA-1 receptor. Ligands for the LDL receptor include modified lipoproteins such as IDL and LDL. Although the LDL receptor is important in lipid metabolism in the liver and steroidogenic tissues, it is not expressed in atherosclerotic lesions. The VLDL receptor specifically bind apoE-containing VLDL and β-VLDL particles as well as Lp(a). The VLDL receptor is expressed in both endothelial and medial SMCs in normal arteries and is also expressed in macrophages in atherosclerotic arteries. LRP mediates uptake of LPL/apoE lipoprotein complex, apoE-enriched VLDL remnants, LPL, LPL-triglyceride-rich lipoprotein complexes, α2-macroglobulin-protease and other protease-antiprotease complexes. LRP is expressed in SMCs and macrophages found in both normal and atherosclerotic lesions. Neither LRP-2

nor apoER2 are expressed in arterial walls, thus these proteins are probably not directly involved in atherogenesis. However, these receptors may contribute to changes in the levels of various lipoproteins in the plasma, thus indirectly promoting artherogenesis. On the other hand, preliminary reports indicate that LR11 is expressed in SMCs of atherosclerotic arteries [Hiltunen et al. (1998) Atherosclerosis 137 Suppl: S81-S88]

Scavenger receptors are expressed on macrophages and specific endothelial cells and mediate the uptake and degradation of polyanionic ligands including modified LDL. Based on structural differences, these receptors are further divided into five classes. Class A scavenger receptors consist of SR-A which encodes three different isoforms (SR-AI, SR-AII, and SR-AIII) due to alternative splicing, and MARCO (macrophage receptor with collagenous structure), all of which bind acetylated LDL. SR-AI and SR-AII receptors are predominantly expressed in macrophages found in atherosclerotic lesions. The Class B scavenger receptors include CD36, SR-BI, an alternatively spliced form of SR-BI designated SR-BII, and the Drosophila *croquemort*. CD36 is expressed on platelets, macrophages, adipocytes, and specific endothelial cells. CD36 binds thrombospondin, collagen, anionic phospholipids, and oxidized LDL among others. SR-BI specifically binds HDL and is able to selectively uptake lipid from HDL thereby removing cholesterol from HDL. SR-BII also functions as an HDL receptor however, it is considerably less efficient in mediating cholesterol transport as compared to SR-BI. The Drosophila dSR-CI, which mediates acetylated LDL uptake by embryonic hemocytes/macrophages, is the only member of the class C scavenger receptors. Class D members include the murine macrosialin and its human homologue CD28. Both bind oxidized LDL and reside in the late endosomal compartment of monocytes and macrophages. Due to their intracellular location, it is speculated that these proteins function in the retention of modified LDL within the cell. The lectin-like oxidized LDL receptor (LOX-1) receptor expressed on endothelial cells defines the class E scavenger receptors and has been shown to preferentially bind oxidized LDL. Finally, class F consists of the scavenger receptor expressed by endothelial cells (SREC) which preferentially binds acetylated LDL. Experiments using knockout mice have verified a role for SR-A as well as other scavenger receptors in the development of atherosclerotic lesions [Greaves et al. (1998) Curr Opin Lipidol 9:425-432].

## 6.1. RESEARCH USES AND UTILITIES

The polynucleotides provided by the present invention can be used by the research community for various purposes. The polynucleotides can be used to express recombinant protein for analysis, characterization or therapeutic use; as markers for tissues in which the corresponding protein is preferentially expressed (either constitutively or at a particular stage of tissue differentiation or development or in disease states); as molecular weight markers on e.g. Southern gels; as chromosome markers or tags (when labeled) to identify chromosomes or to map related gene positions; to compare with endogenous DNA sequences in patients to identify potential genetic disorders; as probes to hybridize and thus discover novel, related DNA sequences; as a source of information to derive PCR primers for genetic fingerprinting; as a probe to "subtract-out" known sequences in the process of discovering other novel polynucleotides; for selecting and making oligomers for attachment to a "gene chip" or other support, including for examination of expression patterns; to raise anti-protein antibodies using DNA immunization techniques; and as an antigen to raise anti-DNA antibodies or elicit another immune response. Where the polynucleotide encodes a protein which binds or potentially binds to another protein (such as, for example, in a receptor-ligand interaction), the polynucleotide can also be used in interaction trap assays (such as, for example, that described in Gyuris et al., Cell 75:791-803 (1993)) to identify polynucleotides encoding the other protein with which binding occurs or to identify inhibitors of the binding interaction.

The proteins provided by the present invention can similarly be used in assays to determine biological activity, including in a panel of multiple proteins for high-throughput screening; to raise antibodies or to elicit another immune response; as a reagent (including the labeled reagent) in assays designed to quantitatively determine levels of the protein (or its receptor) in biological fluids; as markers for tissues in which the corresponding protein is preferentially expressed (either constitutively or at a particular stage of tissue differentiation or development or in a disease state); and, of course, to isolate correlative receptors or ligands. Where the protein binds or potentially binds to another protein (such as, for example, in a receptor-ligand interaction), the protein can be used to identify the other protein with which binding occurs or to identify inhibitors of the binding interaction. Proteins involved in these binding interactions can

also be used to screen for peptide or small molecule inhibitors or agonists of the binding interaction.

Any or all of these research utilities are capable of being developed into reagent grade or kit format for commercialization as research products.

Methods for performing the uses listed above are well known to those skilled in the art. References disclosing such methods include without limitation "Molecular Cloning: A Laboratory Manual", 2d ed., Cold Spring Harbor Laboratory Press, Sambrook, J., E. F. Fritsch and T. Maniatis eds., 1989, and "Methods in Enzymology: Guide to Molecular Cloning Techniques", Academic Press, Berger, S. L. and A. R. Kimmel eds., 1987.

### 6.2. NUTRITIONAL USES

Polynucleotides and proteins of the present invention can also be used as nutritional sources or supplements. Such uses include without limitation use as a protein or amino acid supplement, use as a carbon source, use as a nitrogen source and use as a source of carbohydrate. In such cases the protein or polynucleotide of the invention can be added to the feed of a particular organism or can be administered as a separate solid or liquid preparation, such as in the form of powder, pills, solutions, suspensions or capsules. In the case of microorganisms, the protein or polynucleotide of the invention can be added to the medium in or on which the microorganism is cultured.

### 6.3. CYTOKINE AND CELL PROLIFERATION/DIFFERENTIATION ACTIVITY

A protein of the present invention may exhibit receptor signaling activity relating to cytokine, cell proliferation (either inducing or inhibiting) or cell differentiation (either inducing or inhibiting) activity or may induce production of other cytokines in certain cell populations. A polynucleotide of the invention can encode a polypeptide exhibiting such attributes. Many protein factors discovered to date, including all known cytokines, have exhibited activity in one or more factor-dependent cell proliferation assays, and hence the assays serve as a convenient confirmation of cytokine activity. The activity of therapeutic compositions of the present invention is evidenced by any one of a number of routine factor dependent cell proliferation assays for cell lines including, without limitation, 32D,

DA2, DA1G, T10, B9, B9/11, BaF3, MC9/G, M+(preB M+), 2E8, RB5, DA1, 123, T1165, HT2, CTLL2, TF-1, Mo7e, CMK, HUVEC, and Caco. Therapeutic compositions of the invention can be used in the following:

Assays for T-cell or thymocyte proliferation include without limitation those described in: Current Protocols in Immunology, Ed by J. E. Coligan, A. M. Kruisbeek, D. H. Margulies, E. M. Shevach, W. Strober, Pub. Greene Publishing Associates and Wiley-Interscience (Chapter 3, *In Vitro* assays for Mouse Lymphocyte Function 3.1-3.19; Chapter 7, Immunologic studies in Humans); Takai et al., J. Immunol. 137:3494-3500, 1986; Bertagnolli et al., J. Immunol. 145:1706-1712, 1990; Bertagnolli et al., Cellular Immunology 133:327-341, 1991; Bertagnolli, et al., I. Immunol. 149:3778-3783, 1992; Bowman et al., I. Immunol. 152:1756-1761, 1994.

Assays for cytokine production and/or proliferation of spleen cells, lymph node cells or thymocytes include, without limitation, those described in: Polyclonal T cell stimulation, Kruisbeek, A. M. and Shevach, E. M. In Current Protocols in Immunology. J. E. e.a. Coligan eds. Vol 1 pp. 3.12.1-3.12.14, John Wiley and Sons, Toronto. 1994; and Measurement of mouse and human interleukin-γ, Schreiber, R. D. In Current Protocols in Immunology. J. E. e.a. Coligan eds. Vol 1 pp. 6.8.1-6.8.8, John Wiley and Sons, Toronto. 1994.

Assays for proliferation and differentiation of hematopoietic and lymphopoietic cells include, without limitation, those described in: Measurement of Human and Murine Interleukin 2 and Interleukin 4, Bottomly, K., Davis, L. S. and Lipsky, P. E. In Current Protocols in Immunology. J. E. e.a. Coligan eds. Vol 1 pp. 6.3.1-6.3.12, John Wiley and Sons, Toronto. 1991; deVries et al., J. Exp. Med. 173:1205-1211, 1991; Moreau et al., Nature 336:690-692, 1988; Greenberger et al., Proc. Natl. Acad. Sci. U.S.A. 80:2931-2938, 1983; Measurement of mouse and human interleukin 6--Nordan, R. In Current Protocols in Immunology. J. E. e.a. Coligan eds. Vol 1 pp. 6.6.1-6.6.5, John Wiley and Sons, Toronto. 1991; Smith et al., Proc. Natl. Aced. Sci. U.S.A. 83:1857-1861, 1986; Measurement of human Interleukin 11--Bennett, F., Giannotti, J., Clark, S. C. and Turner, K. J. In Current Protocols in Immunology. J. E. e.a. Coligan eds. Vol 1 pp. 6.15.1 John Wiley and Sons, Toronto. 1991; Measurement of mouse and human Interleukin 9--Ciarletta, A., Giannotti, J., Clark, S. C. and Turner, K. J. In Current Protocols in

Immunology. J. E. e.a. Coligan eds. Vol 1 pp. 6.13.1, John Wiley and Sons, Toronto. 1991.

Assays for T-cell clone responses to antigens (which will identify, among others, proteins that affect APC-T cell interactions as well as direct T-cell effects by measuring proliferation and cytokine production) include, without limitation, those described in: Current Protocols in Immunology, Ed by J. E. Coligan, A. M. Kruisbeek, D. H. Margulies, E. M. Shevach, W Strober, Pub. Greene Publishing Associates and Wiley-Interscience (Chapter 3, *In Vitro* assays for Mouse Lymphocyte Function; Chapter 6, Cytokines and their cellular receptors; Chapter 7, Immunologic studies in Humans); Weinberger et al., Proc. Natl. Acad. Sci. USA 77:6091-6095, 1980; Weinberger et al., Eur. J. Immun. 11:405-411, 1981; Takai et al., J. Immunol. 137:3494-3500, 1986; Takai et al., J. Immunol. 140:508-512, 1988.

### 6.4. IMMUNE STIMULATING OR SUPPRESSING ACTIVITY

A protein of the present invention may also exhibit immune stimulating or immune suppressing activity, including without limitation the activities for which assays are described herein. A polynucleotide of the invention can encode a polypeptide involved in such activities. A protein or antibody, other binding partner, or other modulator of the invention may be useful in the treatment of various immune deficiencies and disorders (including severe combined immunodeficiency (SCID)), e.g., in regulating (up or down) growth and proliferation of T and/or B lymphocytes, as well as effecting the cytolytic activity of NK cells and other cell populations. These immune deficiencies may be genetic or be caused by viral (e.g., HIV) as well as bacterial or fungal infections, or may result from autoimmune disorders. More specifically, infectious diseases caused by viral, bacterial, fungal or other infection may be treatable using a protein, antibody, binding partner, or other modulator of the invention, including infections by HIV, hepatitis viruses, herpesviruses, mycobacteria, Leishmania spp., malaria spp. and various fungal infections such as candidiasis, as well as other conditions where a boost to the immune system generally may be desirable, e.g., in the treatment of cancer.

Autoimmune disorders which may involve a receptor protein of the present invention include, for example, connective tissue disease, multiple sclerosis, systemic lupus erythematosus, rheumatoid arthritis, autoimmune pulmonary inflammation,

Guillain-Barre syndrome, autoimmune thyroiditis, insulin dependent diabetes mellitis, myasthenia gravis, graft-versus-host disease and autoimmune inflammatory eye disease. Such a receptor protein of the present invention may also to be involved in allergic reactions and conditions, such as asthma (particularly allergic asthma) or other respiratory problems.

Using the proteins, antibody, binding partners, or other modulators of the invention it may also be possible to modulate immune responses, in a number of ways. The immune response may be enhanced or suppressed. Down regulation may be in the form of inhibiting or blocking an immune response already in progress or may involve preventing the induction of an immune response. The functions of activated T cells may be inhibited by suppressing T cell responses or by inducing specific tolerance in T cells, or both. Immunosuppression of T cell responses is generally an active, non-antigen-specific, process which requires continuous exposure of the T cells to the suppressive agent. Tolerance, which involves inducing non-responsiveness or anergy in T cells, is distinguishable from immunosuppression in that it is generally antigen-specific and persists after exposure to the tolerizing agent has ceased. Operationally, tolerance can be demonstrated by the lack of a T cell response upon reexposure to specific antigen in the absence of the tolerizing agent.

Down regulating or preventing the immune response, e.g., preventing high level lymphokine synthesis by activated T cells, will be useful in situations of tissue, skin and organ transplantation and in graft-versus-host disease (GVHD). For example, blockage of T cell function should result in reduced tissue destruction in tissue transplantation. Typically, in tissue transplants, rejection of the transplant is initiated through its recognition as foreign by T cells, followed by an immune reaction that destroys the transplant. The administration of a molecule which inhibits or blocks the immune response (e.g. a receptor fragment, binding partner, or other modulator such as antisense polynucleotides) may act as an immunosuppressant.

The efficacy of particular immune response modulators in preventing organ transplant rejection or GVHD can be assessed using animal models that are predictive of efficacy in humans. Examples of appropriate systems which can be used include allogeneic cardiac grafts in rats and xenogeneic pancreatic islet cell grafts in mice, both of which have been used to examine the immunosuppressive effects of CTLA4Ig fusion

proteins *in vivo* as described in Lenschow et al., Science 257:789-792 (1992) and Turka et al., Proc. Natl. Acad. Sci USA, 89:11102-11105 (1992). In addition, murine models of GVHD (see Paul ed., Fundamental Immunology, Raven Press, New York, 1989, pp. 846-847) can be used to determine the effect of blocking B lymphocyte antigen function *in vivo* on the development of that disease.

Blocking the inflammatory response may also be therapeutically useful for treating autoimmune diseases. Many autoimmune disorders are the result of inappropriate activation of T cells that are reactive against self tissue and which promote the production of cytokines and autoantibodies involved in the pathology of the diseases. Preventing the activation of autoreactive T cells may reduce or eliminate disease symptoms. Administration of reagents which block costimulation of T cells can be used to inhibit T cell activation and prevent production of autoantibodies or T cell-derived cytokines which may be involved in the disease process. Additionally, blocking reagents may induce antigen-specific tolerance of autoreactive T cells which could lead to long-term relief from the disease. The efficacy of blocking reagents in preventing or alleviating autoimmune disorders can be determined using a number of well-characterized animal models of human autoimmune diseases. Examples include murine experimental autoimmune encephalitis, systemic lupus erythmatosis in MRL/lpr/lpr mice or NZB hybrid mice, murine autoimmune collagen arthritis, diabetes mellitus in NOD mice and BB rats, and murine experimental myasthenia gravis (see Paul ed., Fundamental Immunology, Raven Press, New York, 1989, pp. 840-856).

Upregulation of immune responses, may also be useful in therapy. Upregulation of immune responses may be in the form of enhancing an existing immune response or eliciting an initial immune response. For example, enhancing an immune response may be useful in cases of viral infection such as influenza, the common cold, and encephalitis.

Alternatively, anti-viral immune responses may be enhanced in an infected patient by removing T cells from the patient, costimulating the T cells *in vitro* and reintroducing the *in vitro* activated T cells into the patient.

The activity of therapeutic compositions of the invention may, among other means, be measured by the following methods:

Suitable assays for thymocyte or splenocyte cytotoxicity include, without limitation, those described in: Current Protocols in Immunology, Ed by J. E. Coligan, A.

M. Kruisbeek, D. H. Margulies, E. M. Shevach, W. Strober, Pub. Greene Publishing Associates and Wiley-Interscience (Chapter 3, *In Vitro* assays for Mouse Lymphocyte Function 3.1-3.19; Chapter 7, Immunologic studies in Humans); Herrmann et al., Proc. Natl. Acad. Sci. USA 78:2488-2492, 1981; Herrmann et al., J. Immunol. 128:1968-1974, 1982; Handa et al., J. Immunol. 135:1564-1572, 1985; Takai et al., I. Immunol. 137:3494-3500, 1986; Takai et al., J. Immunol. 140:508-512, 1988; Herrmann et al., Proc. Natl. Acad. Sci. USA 78:2488-2492, 1981; Herrmann et al., J. Immunol. 128:1968-1974, 1982; Handa et al., J. Immunol. 135:1564-1572, 1985; Takai et al., J. Immunol. 137:3494-3500, 1986; Bowmanet al., J. Virology 61:1992-1998; Takai et al., J. Immunol. 140:508-512, 1988; Bertagnolli et al., Cellular Immunology 133:327-341, 1991; Brown et al., J. Immunol. 153:3079-3092, 1994.

Assays for T-cell-dependent immunoglobulin responses and isotype switching (which will identify, among others, proteins that modulate T-cell dependent antibody responses and that affect Th1/Th2 profiles) include, without limitation, those described in: Maliszewski, J. Immunol. 144:3028-3033, 1990; and Assays for B cell function: *In vitro* antibody production, Mond, J. J. and Brunswick, M. In Current Protocols in Immunology. J. E. e.a. Coligan eds. Vol 1 pp. 3.8.1-3.8.16, John Wiley and Sons, Toronto. 1994.

Mixed lymphocyte reaction (MLR) assays (which will identify, among others, proteins that generate predominantly Th1 and CTL responses) include, without limitation, those described in: Current Protocols in Immunology, Ed by J. E. Coligan, A. M. Kruisbeek, D. H. Margulies, E. M. Shevach, W. Strober, Pub. Greene Publishing Associates and Wiley-Interscience (Chapter 3, *In Vitro* assays for Mouse Lymphocyte Function 3.1-3.19; Chapter 7, Immunologic studies in Humans); Takai et al., J. Immunol. 137:3494-3500, 1986; Takai et al., J. Immunol. 140:508-512, 1988; Bertagnolli et al., J. Immunol. 149:3778-3783, 1992.

Dendritic cell-dependent assays (which will identify, among others, proteins expressed by dendritic cells that activate naive T-cells) include, without limitation, those described in: Guery et al., J. Immunol. 134:536-544, 1995; Inaba et al., Journal of Experimental Medicine 173:549-559, 1991; Macatonia et al., Journal of Immunology 154:5071-5079, 1995; Porgador et al., Journal of Experimental Medicine 182:255-260, 1995; Nair et al., Journal of Virology 67:4062-4069, 1993; Huang et al., Science

264:961-965, 1994; Macatonia et al., Journal of Experimental Medicine 169:1255-1264, 1989; Bhardwaj et al., Journal of Clinical Investigation 94:797-807, 1994; and Inaba et al., Journal of Experimental Medicine 172:631-640, 1990.

Assays for lymphocyte survival/apoptosis (which will identify, among others, proteins that prevent apoptosis after superantigen induction and proteins that regulate lymphocyte homeostasis) include, without limitation, those described in: Darzynkiewicz et al., Cytometry 13:795-808, 1992; Gorczyca et al., Leukemia 7:659-670, 1993; Gorczyca et al., Cancer Research 53:1945-1951, 1993; Itoh et al., Cell 66:233-243, 1991; Zacharchuk, Journal of Immunology 145:4037-4045, 1990; Zamai et al., Cytometry 14:891-897, 1993; Gorczyca et al., International Journal of Oncology 1:639-648, 1992.

Assays for proteins that influence early steps of T-cell commitment and development include, without limitation, those described in: Antica et al., Blood 84:111-117, 1994; Fine et al., Cellular Immunology 155:111-122, 1994; Galy et al., Blood 85:2770-2778, 1995; Toki et al., Proc. Nat. Acad Sci. USA 88:7548-7551, 1991.

## 6.5. HEMATOPOIESIS REGULATING ACTIVITY

A protein of the present invention may be involved in regulation of hematopoiesis and, consequently, in the treatment of myeloid or lymphoid cell deficiencies. Even marginal biological activity in support of colony forming cells or of factor-dependent cell lines indicates involvement in regulating hematopoiesis, e.g. in supporting the growth and proliferation of erythroid progenitor cells alone or in combination with other cytokines, thereby indicating utility, for example, in treating various anemias or for use in conjunction with irradiation/chemotherapy to stimulate the production of erythroid precursors and/or erythroid cells; in supporting the growth and proliferation of myeloid cells such as granulocytes and monocytes/macrophages (i.e., traditional CSF activity) useful, for example, in conjunction with chemotherapy to prevent or treat consequent myelo-suppression; in supporting the growth and proliferation of megakaryocytes and consequently of platelets thereby allowing prevention or treatment of various platelet disorders such as thrombocytopenia, and generally for use in place of or complimentary to platelet transfusions; and/or in supporting the growth and proliferation of hematopoietic stem cells which are capable of maturing to any and all of the above-mentioned hematopoietic cells and therefore find therapeutic utility in various stem cell disorders

(such as those usually treated with transplantation, including, without limitation, aplastic anemia and paroxysmal nocturnal hemoglobinuria), as well as in repopulating the stem cell compartment post irradiation/chemotherapy, either in-vivo or ex-vivo (i.e., in conjunction with bone marrow transplantation or with peripheral progenitor cell transplantation (homologous or heterologous)) as normal cells or genetically manipulated for gene therapy.

Therapeutic compositions of the invention can be used in the following:

Suitable assays for proliferation and differentiation of various hematopoietic lines are cited above.

Assays for embryonic stem cell differentiation (which will identify, among others, proteins that influence embryonic differentiation hematopoiesis) include, without limitation, those described in: Johansson et al. Cellular Biology 15:141-151, 1995; Keller et al., Molecular and Cellular Biology 13:473-486, 1993; McClanahan et al., Blood 81:2903-2915, 1993.

Assays for stem cell survival and differentiation (which will identify, among others, proteins that regulate lympho-hematopoiesis) include, without limitation, those described in: Methylcellulose colony forming assays, Freshney, M. G. In Culture of Hematopoietic Cells. R. I. Freshney, et al. eds. Vol pp. 265-268, Wiley-Liss, Inc., New York, N.Y. 1994; Hirayama et al., Proc. Natl. Acad. Sci. USA 89:5907-5911, 1992; Primitive hematopoietic colony forming cells with high proliferative potential, McNiece, I. K. and Briddell, R. A. In Culture of Hematopoietic Cells. R. I. Freshney, et al. eds. Vol pp. 23-39, Wiley-Liss, Inc., New York, N.Y. 1994; Neben et al., Experimental Hematology 22:353-359, 1994; Cobblestone area forming cell assay, Ploemacher, R. E. In Culture of Hematopoietic Cells. R. I. Freshney, et al. eds. Vol pp. 1-21, Wiley-Liss, Inc., New York, N.Y. 1994; Long term bone marrow cultures in the presence of stromal cells, Spooncer, E., Dexter, M. and Allen, T. In Culture of Hematopoietic Cells. R. I. Freshney, et al. eds. Vol pp. 163-179, Wiley-Liss, Inc., New York, N.Y. 1994; Long term culture initiating cell assay, Sutherland, H. J. In Culture of Hematopoietic Cells. R. I. Freshney, et al. eds. Vol pp. 139-162, Wiley-Liss, Inc., New York, N.Y. 1994.

## 6.6. TISSUE GROWTH ACTIVITY

A protein of the present invention also may be involved in bone, cartilage, tendon, ligament and/or nerve tissue growth or regeneration, as well as in wound healing and tissue repair and replacement, and in healing of burns, incisions and ulcers.

For example, induction of cartilage and/or bone growth in circumstances where bone is not normally formed, has application in the healing of bone fractures and cartilage damage or defects in humans and other animals. Compositions of a protein, antibody, binding partner, or other modulator of the invention may have prophylactic use in closed as well as open fracture reduction and also in the improved fixation of artificial joints. De novo bone formation induced by an osteogenic agent contributes to the repair of congenital, trauma induced, or oncologic resection induced craniofacial defects, and also is useful in cosmetic plastic surgery.

A protein of this invention may also be involved in attracting bone-forming cells, stimulating growth of bone-forming cells, or inducing differentiation of progenitors of bone-forming cells. Treatment of osteoporosis, osteoarthritis, bone degenerative disorders, or periodontal disease, such as through stimulation of bone and/or cartilage repair or by blocking inflammation or processes of tissue destruction (collagenase activity, osteoclast activity, etc.) mediated by inflammatory processes may also be possible using the composition of the invention.

Another category of tissue regeneration activity that may involve the protein of the present invention is tendon/ligament formation. Induction of tendon/ligament-like tissue or other tissue formation in circumstances where such tissue is not normally formed, has application in the healing of tendon or ligament tears, deformities and other tendon or ligament defects in humans and other animals. Such a preparation employing a tendon/ligament-like tissue inducing protein may have prophylactic use in preventing damage to tendon or ligament tissue, as well as use in the improved fixation of tendon or ligament to bone or other tissues, and in repairing defects to tendon or ligament tissue. De novo tendon/ligament-like tissue formation induced by a composition of the present invention contributes to the repair of congenital, trauma induced, or other tendon or ligament defects of other origin, and is also useful in cosmetic plastic surgery for attachment or repair of tendons or ligaments. The compositions of the present invention may provide environment to attract tendon- or ligament-forming cells, stimulate growth of tendon- or ligament-forming cells, induce differentiation of progenitors of tendon- or

ligament-forming cells, or induce growth of tendon/ligament cells or progenitors *ex vivo* for return *in vivo* to effect tissue repair. The compositions of the invention may also be useful in the treatment of tendinitis, carpal tunnel syndrome and other tendon or ligament defects. The compositions may also include an appropriate matrix and/or sequestering agent as a carrier as is well known in the art.

The compositions of the present invention may also be useful for proliferation of neural cells and for regeneration of nerve and brain tissue, i.e. for the treatment of central and peripheral nervous system diseases and neuropathies, as well as mechanical and traumatic disorders, which involve degeneration, death or trauma to neural cells or nerve tissue. More specifically, a composition may be used in the treatment of diseases of the peripheral nervous system, such as peripheral nerve injuries, peripheral neuropathy and localized neuropathies, and central nervous system diseases, such as Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic lateral sclerosis, and Shy-Drager syndrome. Further conditions which may be treated in accordance with the present invention include mechanical and traumatic disorders, such as spinal cord disorders, head trauma and cerebrovascular diseases such as stroke. Peripheral neuropathies resulting from chemotherapy or other medical therapies may also be treatable using a composition of the invention.

Compositions of the invention may also be useful to promote better or faster closure of non-healing wounds, including without limitation pressure ulcers, ulcers associated with vascular insufficiency, surgical and traumatic wounds, and the like.

Compositions of the present invention may also be involved in the generation or regeneration of other tissues, such as organs (including, for example, pancreas, liver, intestine, kidney, skin, endothelium), muscle (smooth, skeletal or cardiac) and vascular (including vascular endothelium) tissue, or for promoting the growth of cells comprising such tissues. Inhibition or modulation of fibrotic scarring may allow normal tissue to regenerate.

A composition of the present invention may also be useful for gut protection or regeneration and treatment of lung or liver fibrosis, reperfusion injury in various tissues, and conditions resulting from systemic cytokine damage.

A composition of the present invention may also be useful for promoting or inhibiting differentiation of tissues described above from precursor tissues or cells; or for inhibiting the growth of tissues described above.

Therapeutic compositions of the invention can be used in the following:

Assays for tissue generation activity include, without limitation, those described in: International Patent Publication No. WO95/16035 (bone, cartilage, tendon); International Patent Publication No. WO95/05846 (nerve, neuronal); International Patent Publication No. WO91/07491 (skin, endothelium).

Assays for wound healing activity include, without limitation, those described in: Winter, Epidermal Wound Healing, pps. 71-112 (Maibach, H. I. and Rovee, D. T., eds.), Year Book Medical Publishers, Inc., Chicago, as modified by Eaglstein and Mertz, J. Invest. Dermatol 71:382-84 (1978).

### 6.7. CHEMOTACTIC/CHEMOKINETIC ACTIVITY

A protein of the present invention may be involved in chemotactic or chemokinetic activity (e.g., act as a chemokine receptor) for mammalian cells, including, for example, monocytes, fibroblasts, neutrophils, T-cells, mast cells, eosinophils, epithelial and/or endothelial cells. A polynucleotide of the invention can encode a polypeptide exhibiting such attributes. Chemotactic and chemokinetic receptor activation can be used to mobilize or attract a desired cell population to a desired site of action. Chemotactic or chemokinetic compositions (e.g. proteins, antibodies, binding partners, or modulators of the invention) provide particular advantages in treatment of wounds and other trauma to tissues, as well as in treatment of localized infections. For example, attraction of lymphocytes, monocytes or neutrophils to tumors or sites of infection may result in improved immune responses against the tumor or infecting agent.

A protein or peptide has chemotactic activity for a particular cell population if it can stimulate, directly or indirectly, the directed orientation or movement of such cell population. Preferably, the protein or peptide has the ability to directly stimulate directed movement of cells. Whether a particular protein has chemotactic activity for a population of cells can be readily determined by employing such protein or peptide in any known assay for cell chemotaxis.

Therapeutic compositions of the invention can be used in the following:

Assays for chemotactic activity (which will identify proteins that induce or prevent chemotaxis) consist of assays that measure the ability of a protein to induce the migration of cells across a membrane as well as the ability of a protein to induce the adhesion of one cell population to another cell population. Suitable assays for movement and adhesion include, without limitation, those described in: Current Protocols in Immunology, Ed by J. E. Coligan, A. M. Kruisbeek, D. H. Marguiles, E. M. Shevach, W. Strober, Pub. Greene Publishing Associates and Wiley-Interscience (Chapter 6.12, Measurement of alpha and beta Chemokines 6.12.1-6.12.28; Taub et al. J. Clin. Invest. 95:1370-1376, 1995; Lind et al. APMIS 103:140-146, 1995; Muller et al Eur. J. Immunol. 25:1744-1748; Gruber et al. J. of Immunol. 152:5860-5867, 1994; Johnston et al. J. of Immunol. 153:1762-1768, 1994.

## 6.8. HEMOSTATIC AND THROMBOLYTIC ACTIVITY

A protein of the invention may also be involved in hemostatis or thrombolysis or thrombosis. A polynucleotide of the invention can encode a polypeptide exhibiting such attributes. Compositions may be useful in treatment of various coagulation disorders (including hereditary disorders, such as hemophilias) or to enhance coagulation and other hemostatic events in treating wounds resulting from trauma, surgery or other causes. A composition of the invention may also be useful for dissolving or inhibiting formation of thromboses and for treatment and prevention of conditions resulting therefrom (such as, for example, infarction of cardiac and central nervous system vessels (e.g., stroke).

Therapeutic compositions of the invention can be used in the following:

Assay for hemostatic and thrombolytic activity include, without limitation, those described in: Linet et al., J. Clin. Pharmacol. 26:131-140, 1986; Burdick et al., Thrombosis Res. 45:413-419, 1987; Humphrey et al., Fibrinolysis 5:71-79 (1991); Schaub, Prostaglandins 35:467-474, 1988.

## 6.10. RECEPTOR/LIGAND ACTIVITY

A protein of the present invention may also demonstrate activity as receptors , receptor ligands or inhibitors or agonists of receptor/ligand interactions. A polynucleotide of the invention can encode a polypeptide exhibiting such characteristics. Examples of such receptors and ligands include, without limitation, cytokine receptors and their

ligands, receptor kinases and their ligands, receptor phosphatases and their ligands, receptors involved in cell-cell interactions and their ligands (including without limitation, cellular adhesion molecules (such as selectins, integrins and their ligands) and receptor/ligand pairs involved in antigen presentation, antigen recognition and development of cellular and humoral immune responses). Receptors and ligands are also useful for screening of potential peptide or small molecule inhibitors of the relevant receptor/ligand interaction. A protein of the present invention (including, without limitation, fragments of receptors and ligands) may themselves be useful as inhibitors of receptor/ligand interactions.

The activity of a protein of the invention may, among other means, be measured by the following methods:

Suitable assays for receptor-ligand activity include without limitation those described in: Current Protocols in Immunology, Ed by J. E. Coligan, A. M. Kruisbeek, D. H. Margulies, E. M. Shevach, W. Strober, Pub. Greene Publishing Associates and Wiley-Interscience (Chapter 7.28, Measurement of Cellular Adhesion under static conditions 7.28.1-7.28.22), Takai et al., Proc. Natl. Acad. Sci. USA 84:6864-6868, 1987; Bierer et al., J. Exp. Med. 168:1145-1156, 1988; Rosenstein et al., J. Exp. Med. 169:149-160 1989; Stoltenborg et al., J. Immunol. Methods 175:59-68, 1994; Stitt et al., Cell 80:661-670, 1995.

By way of example, the CG27, CG153 or CG168 polypeptides of the invention may be used as a lipoprotein receptor for a ligand(s) thereby transmitting the biological activity of that ligand(s). Ligands may be identified through binding assays, affinity chromatography, dihybrid screening assays, BIAcore assays, gel overlay assays, or other methods known in the art.

Studies characterizing drugs or proteins as agonist or antagonist or partial agonists or a partial antagonist require the use of other proteins as competing ligands. The polypeptides of the present invention or ligand(s) thereof may be labeled by being coupled to radioisotopes, colorimetric molecules or a toxin molecules by conventional methods. ("Guide to Protein Purification" Murray P. Deutscher (ed) Methods in Enzymology Vol. 182 (1990) Academic Press, Inc. San Diego). Examples of radioisotopes include, but are not limited to, tritium and carbon-14 . Examples of colorimetric molecules include, but are not limited to, fluorescent molecules such as

fluorescamine, or rhodamine or other colorimetric molecules. Examples of toxins include, but are not limited, to ricin.

### 6.11 DRUG SCREENING

This invention is particularly useful for screening compounds by using the apolipoprotein, lipase or lipoprotein receptor polypeptides of the invention, particularly binding fragments, in any of a variety of drug screening techniques. The polypeptides employed in such a test may either be free in solution, affixed to a solid support, borne on a cell surface or located intracellularly. One method of drug screening utilizes eukaryotic or prokaryotic host cells which are stably transformed with recombinant nucleic acids expressing the desired polypeptide. Drugs are screened against such transformed cells in competitive binding assays. Such cells, either in viable or fixed form, can be used for standard binding assays. One may measure, for example, the formation of complexes between polypeptides of the invention and the agent being tested or examine the diminution in complex formation between the polypeptides and an appropriate cell line, which are well known in the art.

#### 6.11.1 ASSAY FOR RECEPTOR ACTIVITY

The invention also provides methods to detect specific binding of a lipoprotein receptor of the invention to a binding partner polypeptide, or specific binding of an apolipoprotein of the invention to a binding partner polypeptide, in particular a receptor polypeptide. The art provides numerous assays particularly useful for identifying previously unknown binding partners for lipoprotein receptor polypeptides of the invention. For example, expression cloning using mammalian or bacterial cells, or dihybrid screening assays can be used to identify polynucleotides encoding binding partners. As another example, affinity chromatography with the appropriate immobilized polypeptide of the invention can be used to isolate polypeptides that recognize and bind a polypeptide of the invention. Ligands for lipoprotein receptor polypeptides of the invention can also be identified by adding lipoproteins or other exogenous ligands, or cocktails of lipoproteins to two cells populations that are genetically identical except for the expression of the lipoprotein receptor of the invention: one cell population expresses the lipoprotein receptor of the invention whereas the other does not. The response of the

two cell populations to the addition of lipoprotein(s) are then compared. Alternatively, an expression library can be co-expressed with the lipoprotein receptor of the invention in cells and assayed for an autocrine response to identify potential ligand(s). As still another example, BIAcore assays, gel overlay assays, or other methods known in the art can be used to identify binding partner polypeptides.

The role of downstream intracellular signaling molecules in the signaling cascade of the lipoprotein receptor-like CG27, CG153 or CG168 can be determined. For example, a chimeric protein in which the cytoplasmic domain of CG27, CG153 or CG168 is fused to the extracellular portion of a protein, whose ligand has been identified, is produced in a host cell. The cell is then incubated with the ligand specific for the extracellular portion of the chimeric protein, thereby activating the chimeric receptor. Known downstream proteins involved in intracellular signaling can then be assayed for expected modifications i.e. phosphorylation. Other methods known to those in the art can also be used to identify signaling molecules involved in CG27, CG153 or CG168 receptor activity.

## 6.12. ANTI-INFLAMMATORY ACTIVITY

Compositions of the present invention may also exhibit anti-inflammatory activity. The anti-inflammatory activity may be achieved by providing a stimulus to cells involved in the inflammatory response, by inhibiting or promoting cell-cell interactions (such as, for example, cell adhesion), by inhibiting or promoting chemotaxis of cells involved in the inflammatory process, inhibiting or promoting cell extravasation, or by stimulating or suppressing production of other factors which more directly inhibit or promote an inflammatory response. Compositions with such activities can be used to treat inflammatory conditions including chronic or acute conditions), including without limitation intimation associated with infection (such as septic shock, sepsis or systemic inflammatory response syndrome (SIRS)), ischemia-reperfusion injury, endotoxin lethality, arthritis, complement-mediated hyperacute rejection, nephritis, cytokine or chemokine-induced lung injury, inflammatory bowel disease, Crohn's disease or resulting from over production of cytokines such as TNF or IL-1. Compositions of the invention may also be useful to treat anaphylaxis and hypersensitivity to an antigenic substance or material. Compositions of this invention may be utilized to prevent or treat condition

such as, but not limited to, utilized, for example, as part of methods for the prevention and/or treatment of disorders involving sepsis, acute pancreatitis, endotoxin shock, cytokine induced shock, rheumatoid arthritis, chronic inflammatory arthritis, pancreatic cell damage from diabetes mellitus type 1, graft versus host disease, inflammatory bowel disease, inflamation associated with pulmonary disease, other autoimmune disease or inflammatory disease, an antiproliferative agent such as for acute or chronic mylegenous leukemia or in the prevention of premature labor secondary to intrauterine infections.

### 6.13. LEUKEMIAS

Leukemias and related disorders may be treated or prevented by administration of a therapeutic that promotes or inhibits function of the polynucleotides and/or polypeptides of the invention. Such leukemias and related disorders include but are not limited to acute leukemia, acute lymphocytic leukemia, acute myelocytic leukemia, myeloblastic, promyelocytic, myelomonocytic, monotypic, erythroleukemia, chronic leukemia, chronic myelocytic (granulocytic) leukemia and chronic lymphocytic leukemia (for a review of such disorders, see Fishman et al., 1985, Medicine, 2d Ed., J.B. Lippincott Co., Philadelphia).

### 6.14. NERVOUS SYSTEM DISORDERS

Nervous system disorders, involving cell types which can be tested for efficacy of intervention with compounds that modulate the activity of the polynucleotides and/or polypeptides of the invention, and which can be treated upon thus observing an indication of therapeutic utility, include but are not limited to nervous system injuries, and diseases or disorders which result in either a disconnection of axons, a diminution or degeneration of neurons, or demyelination. Nervous system lesions which may be treated in a patient (including human and non-human mammalian patients) according to the invention include but are not limited to the following lesions of either the central (including spinal cord, brain) or peripheral nervous systems:

(i) traumatic lesions, including lesions caused by physical injury or associated with surgery, for example, lesions which sever a portion of the nervous system, or compression injuries;

(ii) ischemic lesions, in which a lack of oxygen in a portion of the nervous system results in neuronal injury or death, including cerebral infarction or ischemia, or spinal cord infarction or ischemia;

(iii) infectious lesions, in which a portion of the nervous system is destroyed or injured as a result of infection, for example, by an abscess or associated with infection by human immunodeficiency virus, herpes zoster, or herpes simplex virus or with Lyme disease, tuberculosis, syphilis;

(iv) degenerative lesions, in which a portion of the nervous system is destroyed or injured as a result of a degenerative process including but not limited to degeneration

associated with Parkinson's disease, Alzheimer's disease, Huntington's chorea, or amyotrophic lateral sclerosis;

(v) lesions associated with nutritional diseases or disorders, in which a portion of the nervous system is destroyed or injured by a nutritional disorder or disorder of metabolism including but not limited to, vitamin B12 deficiency, folic acid deficiency, Wernicke disease, tobacco-alcohol amblyopia, Marchiafava-Bignami disease (primary degeneration of the corpus callosum), and alcoholic cerebellar degeneration;

(vi) neurological lesions associated with systemic diseases including but not limited to diabetes (diabetic neuropathy, Bell's palsy), systemic lupus erythematosus, carcinoma, or sarcoidosis;

(vii) lesions caused by toxic substances including alcohol, lead, or particular neurotoxins; and

(viii) demyelinated lesions in which a portion of the nervous system is destroyed or injured by a demyelinating disease including but not limited to multiple sclerosis, human immunodeficiency virus-associated myelopathy, transverse myelopathy or various etiologies, progressive multifocal leukoencephalopathy, and central pontine myelinolysis.

Therapeutics which are useful according to the invention for treatment of a nervous system disorder may be selected by testing for biological activity in promoting the survival or differentiation of neurons. For example, and not by way of limitation, therapeutics which elicit any of the following effects may be useful according to the invention:

(i) increased survival time of neurons in culture;

(ii) increased sprouting of neurons in culture or *in vivo*;

(iii) increased production of a neuron-associated molecule in culture or *in vivo*, e.g., choline acetyltransferase or acetylcholinesterase with respect to motor neurons; or

(iv) decreased symptoms of neuron dysfunction *in vivo*.

Such effects may be measured by any method known in the art. In preferred, non-limiting embodiments, increased survival of neurons may be measured by the method set forth in Arakawa et al. (1990, J. Neurosci. 10:3507-3515); increased sprouting of neurons may be detected by methods set forth in Pestronk et al. (1980, Exp. Neurol. 70:65-82) or Brown et al. (1981, Ann. Rev. Neurosci. 4:17-42); increased production of neuron-associated molecules may be measured by bioassay, enzymatic assay, antibody binding, Northern

blot assay, etc., depending on the molecule to be measured; and motor neuron dysfunction may be measured by assessing the physical manifestation of motor neuron disorder, e.g., weakness, motor neuron conduction velocity, or functional disability.

In a specific embodiment, motor neuron disorders that may be treated according to the invention include but are not limited to disorders such as infarction, infection, exposure to toxin, trauma, surgical damage, degenerative disease or malignancy that may affect motor neurons as well as other components of the nervous system, as well as disorders that selectively affect neurons such as amyotrophic lateral sclerosis, and including but not limited to progressive spinal muscular atrophy, progressive bulbar palsy, primary lateral sclerosis, infantile and juvenile muscular atrophy, progressive bulbar paralysis of childhood (Fazio-Londe syndrome), poliomyelitis and the post polio syndrome, and Hereditary Motorsensory Neuropathy (Charcot-Marie-Tooth Disease).

## 6.15. OTHER ACTIVITIES

A protein of the invention may also exhibit or be involved in one or more of the following additional activities or effects: inhibiting the growth, infection or function of, or killing, infectious agents, including, without limitation, bacteria, viruses, fungi and other parasites; effecting (suppressing or enhancing) bodily characteristics, including, without limitation, height, weight, hair color, eye color, skin, fat to lean ratio or other tissue pigmentation, or organ or body part size or shape (such as, for example, breast augmentation or diminution, change in bone form or shape); effecting biorhythms or caricadic cycles or rhythms; effecting the fertility of male or female subjects; effecting the metabolism, catabolism, anabolism, processing, utilization, storage or elimination of dietary fat, lipid, protein, carbohydrate, vitamins, minerals, co-factors or other nutritional factors or component(s); effecting behavioral characteristics, including, without limitation, appetite, libido, stress, cognition (including cognitive disorders), depression (including depressive disorders) and violent behaviors; providing analgesic effects or other pain reducing effects; promoting differentiation and growth of embryonic stem cells in lineages other than hematopoietic lineages; hormonal or endocrine activity; in the case of enzymes, correcting deficiencies of the enzyme and treating deficiency-related diseases; treatment of hyperproliferative disorders (such as, for example, psoriasis); immunoglobulin-like activity (such as, for example, the ability to bind antigens or

complement); and the ability to act as an antigen in a vaccine composition to raise an immune response against such protein or another material or entity which is cross-reactive with such protein.

**6.16 IDENTIFICATION OF POLYMORPHISMS**

The demonstration of polymorphisms makes possible the identification of such polymorphisms in human subjects and the pharmacogenetic use of this information for diagnosis and treatment. Such polymorphisms may be associated with, e.g., differential predisposition or susceptibility to various disease states (such as disorders involving inflammation or immune response) or a differential response to drug administration, and this genetic information can be used to tailor preventive or therapeutic treatment appropriately. For example, the existence of a polymorphism associated with a predisposition to inflammation or autoimmune disease makes possible the diagnosis of this condition in humans by identifying the presence of the polymorphism.

Polymorphisms can be identified in a variety of ways known in the art which all generally involve obtaining a sample from a patient, analyzing DNA from the sample, optionally involving isolation or amplification of the DNA, and identifying the presence of the polymorphism in the DNA. For example, PCR may be used to amplify an appropriate fragment of genomic DNA which may then be sequenced. Alternatively, the DNA may be subjected to allele-specific oligonucleotide hybridization (in which appropriate oligonucleotides are hybridized to the DNA under conditions permitting detection of a single base mismatch) or to a single nucleotide extension assay (in which an oligonucleotide that hybridizes immediately adjacent to the position of the polymorphism is extended with one or more labeled nucleotides). In addition, traditional restriction fragment length polymorphism analysis (using restriction enzymes that provide differential digestion of the genomic DNA depending on the presence or absence of the polymorphism) may be performed.

Alternatively a polymorphism resulting in a change in the amino acid sequence could also be detected by detecting a corresponding change in amino acid sequence of the protein, e.g., by an antibody specific to the variant sequence.

## 6.17 CANCER DIAGNOSIS AND THERAPY

Polypeptides of the invention may be involved in cancer cell generation, proliferation or metastasis. Detection of the presence or amount of polynucleotides or polypeptides of the invention may be useful for the diagnosis and/or prognosis of one or more types of cancer. For example, the presence or increased expression of a polynucleotide/polypeptide of the invention may indicate a hereditary risk of cancer, a precancerous condition, or an ongoing malignancy. Conversely, a defect in the gene or absence of the polypeptide may be associated with a cancer condition. Identification of single nucleotide polymorphisms associated with cancer or a predisposition to cancer may also be useful for diagnosis or prognosis.

Cancer treatments promote tumor regression by inhibiting tumor cell proliferation, inhibiting angiogenesis (growth of new blood vessels that is necessary to support tumor growth) and/or prohibiting metastasis by reducing tumor cell motility or invasiveness. Therapeutic compositions of the invention may be effective in adult and pediatric oncology including in solid phase tumors/malignancies, locally advanced tumors, human soft tissue sarcomas, metastatic cancer, including lymphatic metastases, blood cell malignancies including multiple myeloma, acute and chronic leukemias, and lymphomas, head and neck cancers including mouth cancer, larynx cancer and thyroid cancer, lung cancers including small cell carcinoma and non-small cell cancers, breast cancers including small cell carcinoma and ductal carcinoma, gastrointestinal cancers including esophageal cancer, stomach cancer, colon cancer, colorectal cancer and polyps associated with colorectal neoplasia, pancreatic cancers, liver cancer, urologic cancers including bladder cancer and prostate cancer, malignancies of the female genital tract including ovarian carcinoma, uterine (including endometrial) cancers, and solid tumor in the ovarian follicle, kidney cancers including renal cell carcinoma, brain cancers including intrinsic brain tumors, neuroblastoma, astrocytic brain tumors, gliomas, metastatic tumor cell invasion in the central nervous system, bone cancers including osteomas, skin cancers including malignant melanoma, tumor progression of human skin keratinocytes, squamous cell carcinoma, basal cell carcinoma, hemangiopericytoma and Karposi's sarcoma.

Polypeptides, polynucleotides, or modulators of polypeptides of the invention (including inhibitors and stimulators of the biological activity of the polypeptide of the

invention) may be administered to treat cancer. Therapeutic compositions can be administered in therapeutically effective dosages alone or in combination with adjuvant cancer therapy such as surgery, chemotherapy, radiotherapy, thermotherapy, and laser therapy, and may provide a beneficial effect, e.g. reducing tumor size, slowing rate of tumor growth, inhibiting metastasis, or otherwise improving overall clinical condition, without necessarily eradicating the cancer.

The composition can also be administered in therapeutically effective amounts as a portion of an anti-cancer cocktail. An anti-cancer cocktail is a mixture of the polypeptide or modulator of the invention with one or more anti-cancer drugs in addition to a pharmaceutically acceptable carrier for delivery. The use of anti-cancer cocktails as a cancer treatment is routine. Anti-cancer drugs that are well known in the art and can be used as a treatment in combination with the polypeptide or modulator of the invention include: Actinomycin D, Aminoglutethimide, Asparaginase, Bleomycin, Busulfan, Carboplatin, Carmustine, Chlorambucil, Cisplatin (cis-DDP), Cyclophosphamide, Cytarabine HCl (Cytosine arabinoside), Dacarbazine, Dactinomycin, Daunorubicin HCl, Doxorubicin HCl, Estramustine phosphate sodium, Etoposide (V16-213), Floxuridine, 5-Fluorouracil (5-Fu), Flutamide, Hydroxyurea (hydroxycarbamide), Ifosfamide, Interferon Alpha-2a, Interferon Alpha-2b, Leuprolide acetate (LHRH-releasing factor analog), Lomustine, Mechlorethamine HCl (nitrogen mustard), Melphalan, Mercaptopurine, Mesna, Methotrexate (MTX), Mitomycin, Mitoxantrone HCl, Octreotide, Plicamycin, Procarbazine HCl, Streptozocin, Tamoxifen citrate, Thioguanine, Thiotepa, Vinblastine sulfate, Vincristine sulfate, Amsacrine, Azacitidine, Hexamethylmelamine, Interleukin-2, Mitoguazone, Pentostatin, Semustine, Teniposide, and Vindesine sulfate.

In addition, therapeutic compositions of the invention may be used for prophylactic treatment of cancer. There are hereditary conditions and/or environmental situations (e.g. exposure to carcinogens) known in the art that predispose an individual to developing cancers. Under these circumstances, it may be beneficial to treat these individuals with therapeutically effective doses of the polypeptide of the invention to reduce the risk of developing cancers.

In vitro models can be used to determine the effective doses of the polypeptide of the invention as a potential cancer treatment. These in vitro models include proliferation assays of cultured tumor cells, growth of cultured tumor cells in soft agar (see Freshney,

(1987) Culture of Animal Cells: A Manual of Basic Technique, Wily-Liss, New York, NY Ch 18 and Ch 21), tumor systems in nude mice as described in Giovanella et al., J. Natl. Can. Inst., 52: 921-30 (1974), mobility and invasive potential of tumor cells in Boyden Chamber assays as described in Pilkington et al., Anticancer Res., 17: 4107-9 (1997), and angiogenesis assays such as induction of vascularization of the chick chorioallantoic membrane or induction of vascular endothelial cell migration as described in Ribatta et al., Intl. J. Dev. Biol., 40: 1189-97 (1999) and Li et al., Clin. Exp. Metastasis, 17:423-9 (1999) respectively. Suitable tumor cells lines are available, e.g. from American Type Tissue Culture Collection catalogs.

## 7. THERAPEUTIC METHODS

The compositions (including polypeptide fragments, analogs, variants and antibodies or other binding partners or modulators including antisense polynucleotides) of the invention have numerous applications in a variety of therapeutic methods. Examples of therapeutic applications include, but are not limited to, those exemplified below.

### 7.1 SEPSIS

One embodiment of the invention is the administration of an effective amount of compositions of the invention to individuals that are at a high risk of developing sepsis, or that have developed sepsis. An example of the former category are patients about to undergo surgery. While the mode of administration is not particularly important, parenteral administration is preferred because of the rapid progression of sepsis, and thus, the need to have the inhibitor disseminate quickly throughout the body. Thus, the preferred mode of administration is to deliver an I.V. bolus slightly before, during, or after surgery. The dosage of the compositions of the invention will normally be determined by the prescribing physician. It is to be expected that the dosage will vary according to the age, weight and response of the individual patient. Typically, where a protein is being administered, the amount of inhibitor administered per dose will be in the range of about 0.1 to 25 mg/kg of body weight, with the preferred dose being about 0.1 to 10 mg/kg of patient body weight. For parenteral administration, the compositions of the invention may be formulated in an injectable form that includes a pharmaceutically acceptable parenteral vehicle. Such vehicles are well known in the art and examples

include water, saline, Ringer's solution, dextrose solution, and solutions consisting of small amounts of the human serum albumin. The vehicle may contain minor amounts of additives that maintain the isotonicity and stability of the inhibitor. The preparation of such solutions is within the skill of the art. Typically, the cytokine inhibitor will be formulated in such vehicles at a concentration of about 1-8 mg/ml to about 10 mg/ml.

## 7.2 ARTHRITIS AND INFLAMMATION

The immunosuppressive effects of the compositions of the invention against rheumatoid arthritis is determined in an experimental animal model system. The experimental model system is adjuvant induced arthritis in rats, and the protocol is described by J. Holoshitz, et at., 1983, Science, 219:56, or by B. Waksman et al., 1963, Int. Arch. Allergy Appl. Immunol., 23:129. Induction of the disease can be caused by a single injection, generally intradermally, of a suspension of killed Mycobacterium tuberculosis in complete Freund's adjuvant (CFA). The route of injection can vary, but rats may be injected at the base of the tail with an adjuvant mixture. The inhibitor is administered in phosphate buffered solution (PBS) at a dose of about 1-5 mg/kg. The control consists of administering PBS only.

The procedure for testing the effects of the test compound would consist of intradermally injecting killed Mycobacterium tuberculosis in CFA followed by immediately administering the inhibitor and subsequent treatment every other day until day 24. At 14, 15, 18, 20, 22, and 24 days after injection of Mycobacterium CFA, an overall arthritis score may be obtained as described by J. Holoskitz above. An analysis of the data would reveal that the test compound would have a dramatic affect on the swelling of the joints as measured by a decrease of the arthritis score.

## 7.4 PHARMACEUTICAL FORMULATIONS AND ROUTES OF ADMINISTRATION

A protein of the present invention (from whatever source derived, including without limitation from recombinant and non-recombinant sources and including antibodies and other binding partners of the polypeptides of the invention) may be administered to a patient in need, by itself, or in pharmaceutical compositions where it is mixed with suitable carriers or excipient(s) at doses to treat or ameliorate a variety of

disorders. Such a composition may also contain (in addition to protein and a carrier) diluents, fillers, salts, buffers, stabilizers, solubilizers, and other materials well known in the art. The term "pharmaceutically acceptable" means a non-toxic material that does not interfere with the effectiveness of the biological activity of the active ingredient(s). The characteristics of the carrier will depend on the route of administration. The pharmaceutical composition of the invention may also contain cytokines, lymphokines, or other hematopoietic factors such as M-CSF, GM-CSF, TNF, IL-1, IL-2, IL-3, IL-4, IL-5, IL-6, IL-7, IL-8, IL-9, IL-10, IL-11, IL-12, IL-13, IL-14, IL-15, IL-16, IL-17, IL-18, IFN, TNF0, TNF1, TNF2, G-CSF, Meg-CSF, GM-CSF, thrombopoietin, stem cell factor, and erythropoietin. In further compositions, proteins of the invention may be combined with other agents beneficial to the treatment of the bone and/or cartilage defect, wound, or tissue in questions. These agents include various growth factors such as epidermal growth factor (EGF), platelet-derived growth factor (PDGF), transforming growth factors (TGF-α and TGF-β), insulin-like growth factor (IGF), as well as cytokines described herein.

The pharmaceutical composition may further contain other agents which either enhance the activity of the protein or compliment its activity or use in treatment. Such additional factors and/or agents may be included in the pharmaceutical composition to produce a synergistic effect with protein of the invention, or to minimize side effects. Conversely, protein of the present invention may be included in formulations of the particular cytokine, lymphokine, other hematopoietic factor, thrombolytic or anti-thrombotic factor, or anti-inflammatory agent to minimize side effects of the cytokine, lymphokine, other hematopoietic factor, thrombolytic or anti-thrombotic factor, or anti-inflammatory agent. A protein of the present invention may be active in multimers (e.g., heterodimers or homodimers) or complexes with itself or other proteins. As a result, pharmaceutical compositions of the invention may comprise a protein of the invention in such multimeric or complexed form.

As an alternative to being included in a pharmaceutical composition of the invention including a first protein, a second protein or a therapeutic agent may be concurrently administered with the first protein.

Techniques for formulation and administration of the compounds of the instant application may be found in "Remington's Pharmaceutical Sciences," Mack Publishing

Co., Easton, PA, latest edition. A therapeutically effective dose further refers to that amount of the compound sufficient to result in amelioration of symptoms, e.g., treatment, healing, prevention or amelioration of the relevant medical condition, or an increase in rate of treatment, healing, prevention or amelioration of such conditions. When applied to an individual active ingredient, administered alone, a therapeutically effective dose refers to that ingredient alone. When applied to a combination, a therapeutically effective dose refers to combined amounts of the active ingredients that result in the therapeutic effect, whether administered in combination, serially or simultaneously.

In practicing the method of treatment or use of the present invention, a therapeutically effective amount of protein of the present invention is administered to a mammal having a condition to be treated. Protein of the present invention may be administered in accordance with the method of the invention either alone or in combination with other therapies such as treatments employing cytokines, lymphokines or other hematopoietic factors. When co-administered with one or more cytokines, lymphokines or other hematopoietic factors, protein of the present invention may be administered either simultaneously with the cytokine(s), lymphokine(s), other hematopoietic factor(s), thrombolytic or anti-thrombotic factors, or sequentially. If administered sequentially, the attending physician will decide on the appropriate sequence of administering protein of the present invention in combination with cytokine(s), lymphokine(s), other hematopoietic factor(s), thrombolytic or anti-thrombotic factors.

## 7.5. ROUTES OF ADMINISTRATION

Suitable routes of administration may, for example, include oral, rectal, transmucosal, or intestinal administration; parenteral delivery, including intramuscular, subcutaneous, intramedullary injections, as well as intrathecal, direct intraventricular, intravenous, intraperitoneal, intranasal, or intraocular injections. Administration of protein of the present invention used in the pharmaceutical composition or to practice the method of the present invention can be carried out in a variety of conventional ways, such as oral ingestion, inhalation, topical application or cutaneous, subcutaneous, intraperitoneal, parenteral or intravenous injection. Intravenous administration to the patient is preferred.

Alternately, one may administer the compound in a local rather than systemic manner, for example, via injection of the compound directly into a arthritic joints or in fibrotic tissue, often in a depot or sustained release formulation. In order to prevent the scarring process frequently occurring as complication of glaucoma surgery, the compounds may be administered topically, for example, as eye drops. Furthermore, one may administer the drug in a targeted drug delivery system, for example, in a liposome coated with a specific antibody, targeting, for example, arthritic or fibrotic tissue. The liposomes will be targeted to and taken up selectively by the afflicted tissue.

## 7.6. COMPOSITIONS/FORMULATIONS

Pharmaceutical compositions for use in accordance with the present invention thus may be formulated in a conventional manner using one or more physiologically acceptable carriers comprising excipients and auxiliaries which facilitate processing of the active compounds into preparations which can be used pharmaceutically. These pharmaceutical compositions may be manufactured in a manner that is itself known, e.g., by means of conventional mixing, dissolving, granulating, dragee-making, levigating, emulsifying, encapsulating, entrapping or lyophilizing processes. Proper formulation is dependent upon the route of administration chosen. When a therapeutically effective amount of protein of the present invention is administered orally, protein of the present invention will be in the form of a tablet, capsule, powder, solution or elixir. When administered in tablet form, the pharmaceutical composition of the invention may additionally contain a solid carrier such as a gelatin or an adjuvant. The tablet, capsule, and powder contain from about 5 to 95% protein of the present invention, and preferably from about 25 to 90% protein of the present invention. When administered in liquid form, a liquid carrier such as water, petroleum, oils of animal or plant origin such as peanut oil, mineral oil, soybean oil, or sesame oil, or synthetic oils may be added. The liquid form of the pharmaceutical composition may further contain physiological saline solution, dextrose or other saccharide solution, or glycols such as ethylene glycol, propylene glycol or polyethylene glycol. When administered in liquid form, the pharmaceutical composition contains from about 0.5 to 90% by weight of protein of the present invention, and preferably from about 1 to 50% protein of the present invention.

When a therapeutically effective amount of protein of the present invention is administered by intravenous, cutaneous or subcutaneous injection, protein of the present invention will be in the form of a pyrogen-free, parenterally acceptable aqueous solution. The preparation of such parenterally acceptable protein solutions, having due regard to pH, isotonicity, stability, and the like, is within the skill in the art. A preferred pharmaceutical composition for intravenous, cutaneous, or subcutaneous injection should contain, in addition to protein of the present invention, an isotonic vehicle such as Sodium Chloride Injection, Ringer's Injection, Dextrose Injection, Dextrose and Sodium Chloride Injection, Lactated Ringer's Injection, or other vehicle as known in the art. The pharmaceutical composition of the present invention may also contain stabilizers, preservatives, buffers, antioxidants, or other additives known to those of skill in the art. For injection, the agents of the invention may be formulated in aqueous solutions, preferably in physiologically compatible buffers such as Hanks's solution, Ringer's solution, or physiological saline buffer. For transmucosal administration, penetrants appropriate to the barrier to be permeated are used in the formulation. Such penetrants are generally known in the art.

For oral administration, the compounds can be formulated readily by combining the active compounds with pharmaceutically acceptable carriers well known in the art. Such carriers enable the compounds of the invention to be formulated as tablets, pills, dragees, capsules, liquids, gels, syrups, slurries, suspensions and the like, for oral ingestion by a patient to be treated. Pharmaceutical preparations for oral use can be obtained solid excipient, optionally grinding a resulting mixture, and processing the mixture of granules, after adding suitable auxiliaries, if desired, to obtain tablets or dragee cores. Suitable excipients are, in particular, fillers such as sugars, including lactose, sucrose, mannitol, or sorbitol; cellulose preparations such as, for example, maize starch, wheat starch, rice starch, potato starch, gelatin, gum tragacanth, methyl cellulose, hydroxypropylmethyl-cellulose, sodium carboxymethylcellulose, and/or polyvinylpyrrolidone (PVP). If desired, disintegrating agents may be added, such as the cross-linked polyvinyl pyrrolidone, agar, or alginic acid or a salt thereof such as sodium alginate. Dragee cores are provided with suitable coatings. For this purpose, concentrated sugar solutions may be used, which may optionally contain gum arabic, talc, polyvinyl pyrrolidone, carbopol gel, polyethylene glycol, and/or titanium dioxide, lacquer solutions,

and suitable organic solvents or solvent mixtures. Dyestuffs or pigments may be added to the tablets or dragee coatings for identification or to characterize different combinations of active compound doses.

Pharmaceutical preparations which can be used orally include push-fit capsules made of gelatin, as well as soft, sealed capsules made of gelatin and a plasticizer, such as glycerol or sorbitol. The push-fit capsules can contain the active ingredients in admixture with filler such as lactose, binders such as starches, and/or lubricants such as talc or magnesium stearate and, optionally, stabilizers. In soft capsules, the active compounds may be dissolved or suspended in suitable liquids, such as fatty oils, liquid paraffin, or liquid polyethylene glycols. In addition, stabilizers may be added. All formulations for oral administration should be in dosages suitable for such administration. For buccal administration, the compositions may take the form of tablets or lozenges formulated in conventional manner.

For administration by inhalation, the compounds for use according to the present invention are conveniently delivered in the form of an aerosol spray presentation from pressurized packs or a nebuliser, with the use of a suitable propellant, e.g., dichlorodifluoromethane, trichlorofluoromethane, dichlorotetrafluoroethane, carbon dioxide or other suitable gas. In the case of a pressurized aerosol the dosage unit may be determined by providing a valve to deliver a metered amount. Capsules and cartridges of, e.g., gelatin for use in an inhaler or insufflator may be formulated containing a powder mix of the compound and a suitable powder base such as lactose or starch. The compounds may be formulated for parenteral administration by injection, e.g., by bolus injection or continuous infusion. Formulations for injection may be presented in unit dosage form, e.g., in ampules or in multi-dose containers, with an added preservative. The compositions may take such forms as suspensions, solutions or emulsions in oily or aqueous vehicles, and may contain formulatory agents such as suspending, stabilizing and/or dispersing agents.

Pharmaceutical formulations for parenteral administration include aqueous solutions of the active compounds in water-soluble form. Additionally, suspensions of the active compounds may be prepared as appropriate oily injection suspensions. Suitable lipophilic solvents or vehicles include fatty oils such as sesame oil, or synthetic fatty acid esters, such as ethyl oleate or triglycerides, or liposomes. Aqueous injection suspensions

may contain substances which increase the viscosity of the suspension, such as sodium carboxymethyl cellulose, sorbitol, or dextran. Optionally, the suspension may also contain suitable stabilizers or agents which increase the solubility of the compounds to allow for the preparation of highly concentrated solutions. Alternatively, the active ingredient may be in powder form for constitution with a suitable vehicle, e.g., sterile pyrogen-free water, before use.

The compounds may also be formulated in rectal compositions such as suppositories or retention enemas, e.g., containing conventional suppository bases such as cocoa butter or other glycerides. In addition to the formulations described previously, the compounds may also be formulated as a depot preparation. Such long acting formulations may be administered by implantation (for example subcutaneously or intramuscularly) or by intramuscular injection. Thus, for example, the compounds may be formulated with suitable polymeric or hydrophobic materials (for example as an emulsion in an acceptable oil) or ion exchange resins, or as sparingly soluble derivatives, for example, as a sparingly soluble salt.

A pharmaceutical carrier for the hydrophobic compounds of the invention is a cosolvent system comprising benzyl alcohol, a nonpolar surfactant, a water-miscible organic polymer, and an aqueous phase. The cosolvent system may be the VPD co-solvent system. VPD is a solution of 3% w/v benzyl alcohol, 8% w/v of the nonpolar surfactant polysorbate 80, and 65% w/v polyethylene glycol 300, made up to volume in absolute ethanol. The VPD co-solvent system (VPD:5W) consists of VPD diluted 1:1 with a 5% dextrose in water solution. This co-solvent system dissolves hydrophobic compounds well, and itself produces low toxicity upon systemic administration. Naturally, the proportions of a co-solvent system may be varied considerably without destroying its solubility and toxicity characteristics. Furthermore, the identity of the co-solvent components may be varied: for example, other low-toxicity nonpolar surfactants may be used instead of polysorbate 80; the fraction size of polyethylene glycol may be varied; other biocompatible polymers may replace polyethylene glycol, e.g. polyvinyl pyrrolidone; and other sugars or polysaccharides may substitute for dextrose. Alternatively, other delivery systems for hydrophobic pharmaceutical compounds may be employed. Liposomes and emulsions are well known examples of delivery vehicles or carriers for hydrophobic drugs. Certain organic solvents such as dimethylsulfoxide also

may be employed, although usually at the cost of greater toxicity. Additionally, the compounds may be delivered using a sustained-release system, such as semipermeable matrices of solid hydrophobic polymers containing the therapeutic agent. Various types of sustained-release materials have been established and are well known by those skilled in the art. Sustained-release capsules may, depending on their chemical nature, release the compounds for a few weeks up to over 100 days. Depending on the chemical nature and the biological stability of the therapeutic reagent, additional strategies for protein stabilization may be employed.

The pharmaceutical compositions also may comprise suitable solid or gel phase carriers or excipients. Examples of such carriers or excipients include but are not limited to calcium carbonate, calcium phosphate, various sugars, starches, cellulose derivatives, gelatin, and polymers such as polyethylene glycols. Many of the compounds of the invention may be provided as salts with pharmaceutically compatible counterions. Such pharmaceutically acceptable base addition salts are those salts which retain the biological effectiveness and properties of the free acids and which are obtained by reaction with inorganic or organic bases such as sodium hydroxide, magnesium hydroxide, ammonia, trialkylamine, dialkylamine, monoalkylamine, dibasic amino acids, sodium acetate, potassium benzoate, triethanol amine and the like.

The pharmaceutical composition of the invention may be in the form of a complex of the protein(s) of present invention along with protein or peptide antigens. The protein and/or peptide antigen will deliver a stimulatory signal to both B and T lymphocytes. B lymphocytes will respond to antigen through their surface immunoglobulin receptor. T lymphocytes will respond to antigen through the T cell receptor (TCR) following presentation of the antigen by MHC proteins. MHC and structurally related proteins including those encoded by class I and class II MHC genes on host cells will serve to present the peptide antigen(s) to T lymphocytes. The antigen components could also be supplied as purified MHC-peptide complexes alone or with co-stimulatory molecules that can directly signal T cells. Alternatively antibodies able to bind surface immunoglobulin and other molecules on B cells as well as antibodies able to bind the TCR and other molecules on T cells can be combined with the pharmaceutical composition of the invention. The pharmaceutical composition of the invention may be in the form of a liposome in which protein of the present invention is combined, in addition to other

pharmaceutically acceptable carriers, with amphipathic agents such as lipids which exist in aggregated form as micelles, insoluble monolayers, liquid crystals, or lamellar layers in aqueous solution. Suitable lipids for liposomal formulation include, without limitation, monoglycerides, diglycerides, sulfatides, lysolecithin, phospholipids, saponin, bile acids, and the like. Preparation of such liposomal formulations is within the level of skill in the art, as disclosed, for example, in U.S. Patent. Nos. 4,235,871; 4,501,728; 4,837,028; and 4,737,323, all of which are incorporated herein by reference.

The amount of protein of the present invention in the pharmaceutical composition of the present invention will depend upon the nature and severity of the condition being treated, and on the nature of prior treatments which the patient has undergone. Ultimately, the attending physician will decide the amount of protein of the present invention with which to treat each individual patient. Initially, the attending physician will administer low doses of protein of the present invention and observe the patient's response. Larger doses of protein of the present invention may be administered until the optimal therapeutic effect is obtained for the patient, and at that point the dosage is not increased further. It is contemplated that the various pharmaceutical compositions used to practice the method of the present invention should contain about 0.01 μg to about 100 mg (preferably about 0.1 μg to about 10 mg, more preferably about 0.1 μg to about 1 mg) of protein of the present invention per kg body weight. If desired, the therapeutic method includes administering the composition topically, systematically, or locally as an implant or device. When administered, the therapeutic composition for use in this invention is, of course, in a pyrogen-free, physiologically acceptable form. Further, the composition may desirably be encapsulated or injected in a viscous form for delivery to the site of bone, cartilage or tissue damage. Topical administration may be suitable for wound healing and tissue repair. Therapeutically useful agents other than a protein of the invention which may also optionally be included in the composition as described above, may alternatively or additionally, be administered simultaneously or sequentially with the composition in the methods of the invention. Preferably for bone and/or cartilage formation, the composition would include a matrix capable of delivering the protein-containing composition to the site of bone and/or cartilage damage, providing a structure for the developing bone and cartilage and optimally capable of being resorbed into the body.

Such matrices may be formed of materials presently in use for other implanted medical applications.

The choice of matrix material is based on biocompatibility, biodegradability, mechanical properties, cosmetic appearance and interface properties. The particular application of the compositions will define the appropriate formulation. Potential matrices for the compositions may be biodegradable and chemically defined calcium sulfate, tricalcium phosphate, hydroxyapatite, polylactic acid, polyglycolic acid and polyanhydrides. Other potential materials are biodegradable and biologically well-defined, such as bone or dermal collagen. Further matrices are comprised of pure proteins or extracellular matrix components. Other potential matrices are nonbiodegradable and chemically defined, such as sintered hydroxyapatite, bioglass, aluminates, or other ceramics. Matrices may be comprised of combinations of any of the above mentioned types of material, such as polylactic acid and hydroxyapatite or collagen and tricalcium phosphate. The bioceramics may be altered in composition, such as in calcium-aluminate-phosphate and processing to alter pore size, particle size, particle shape, and biodegradability. Presently preferred is a 50:50 (mole weight) copolymer of lactic acid and glycolic acid in the form of porous particles having diameters ranging from 150 to 800 microns. In some applications, it will be useful to utilize a sequestering agent, such as carboxymethyl cellulose or autologous blood clot, to prevent the protein compositions from disassociating from the matrix.

A preferred family of sequestering agents is cellulosic materials such as alkylcelluloses (including hydroxyalkylcelluloses), including methylcellulose, ethylcellulose, hydroxyethylcellulose, hydroxypropylcellulose, hydroxypropyl-methylcellulose, and carboxymethylcellulose, the most preferred being cationic salts of carboxymethylcellulose (CMC). Other preferred sequestering agents include hyaluronic acid, sodium alginate, poly(ethylene glycol), polyoxyethylene oxide, carboxyvinyl polymer and poly(vinyl alcohol). The amount of sequestering agent useful herein is 0.5-20 wt %, preferably 1-10 wt % based on total formulation weight, which represents the amount necessary to prevent desorbtion of the protein from the polymer matrix and to provide appropriate handling of the composition, yet not so much that the progenitor cells are prevented from infiltrating the matrix, thereby providing the protein the opportunity to assist the fracture repair activity of the progenitor cells.

The therapeutic compositions are also presently valuable for veterinary applications. Particularly domestic animals and thoroughbred horses, in addition to humans, are desired patients for such treatment with proteins of the present invention. The dosage regimen of a protein-containing pharmaceutical composition to be used in tissue regeneration will be determined by the attending physician considering various factors which modify the action of the proteins, e.g., amount of tissue weight desired to be formed, the site of damage, the condition of the damaged tissue, the size of a wound, type of damaged tissue (e.g., bone), the patient's age, sex, and diet, the severity of any infection, time of administration and other clinical factors. The dosage may vary with the type of matrix used in the reconstitution and with inclusion of other proteins in the pharmaceutical composition. For example, the addition of other known growth factors, such as IGF I (insulin like growth factor I), to the final composition, may also effect the dosage. Progress can be monitored by periodic assessment of tissue/bone growth and/or repair, for example, X-rays, histomorphometric determinations and tetracycline labeling.

## 7.7. EFFECTIVE DOSAGE

Pharmaceutical compositions suitable for use in the present invention include compositions wherein the active ingredients are contained in an effective amount to achieve its intended purpose. More specifically, a therapeutically effective amount means an amount effective to prevent development of or to alleviate the existing symptoms of the subject being treated. Determination of the effective amounts is well within the capability of those skilled in the art, especially in light of the detailed disclosure provided herein. For any compound used in the method of the invention, the therapeutically effective dose can be estimated initially from appropriate *in vitro* assays. Such information can be used to more accurately determine useful doses in humans.

A therapeutically effective dose refers to that amount of the compound that results in amelioration of symptoms or a prolongation of survival in a patient. Toxicity and therapeutic efficacy of such compounds can be determined by standard pharmaceutical procedures in cell cultures or experimental animals, e.g., for determining the $LD_{50}$ (the dose lethal to 50% of the population) and the $ED_{50}$ (the dose therapeutically effective in 50% of the population). The dose ratio between toxic and therapeutic effects is the therapeutic index and it can be expressed as the ratio between $LD_{50}$ and $ED_{50}$.

Compounds which exhibit high therapeutic indices are preferred. The data obtained from these cell culture assays and animal studies can be used in formulating a range of dosage for use in human. The dosage of such compounds lies preferably within a range of circulating concentrations that include the $ED_{50}$ with little or no toxicity. The dosage may vary within this range depending upon the dosage form employed and the route of administration utilized. The exact formulation, route of administration and dosage can be chosen by the individual physician in view of the patient's condition. See, e.g., Fingl et al., 1975, in "The Pharmacological Basis of Therapeutics", Ch. 1 p.1. Dosage amount and interval may be adjusted individually to provide plasma levels of the active agent which are sufficient to maintain the desired effects, or minimal effective concentration (MEC). The MEC will vary for each compound but can be estimated from *in vitro*. Dosages necessary to achieve the MEC will depend on individual characteristics and route of administration. However, HPLC assays or bioassays can be used to determine plasma concentrations.

Dosage intervals can also be determined using MEC value. Compounds should be administered using a regimen which maintains plasma levels above the MEC for 10-90% of the time, preferably between 30-90% and most preferably between 50-90%. In cases of local administration or selective uptake, the effective local concentration of the drug may not be related to plasma concentration.

An exemplary dosage regimen for the human polypeptides of the invention will be in the range of about 0.01 to 100 mg/kg of body weight daily, with the preferred dose being about 0.1 to 25 mg/kg of patient body weight daily, varying in adults and children. Dosing may be once daily, or equivalent doses may be delivered at longer or shorter intervals.

The amount of composition administered will, of course, be dependent on the subject being treated, on the subject's age and weight, the severity of the affliction, the manner of administration and the judgment of the prescribing physician.

**7.8. PACKAGING**

The compositions may, if desired, be presented in a pack or dispenser device which may contain one or more unit dosage forms containing the active ingredient. The pack may, for example, comprise metal or plastic foil, such as a blister pack. The pack or

dispenser device may be accompanied by instructions for administration. Compositions comprising a compound of the invention formulated in a compatible pharmaceutical carrier may also be prepared, placed in an appropriate container, and labeled for treatment of an indicated condition.

## 8. ANTIBODIES

Another aspect of the invention is an antibody that specifically binds the apolipoprotein, lipase, or lipoprotein receptor polypeptide of the invention. Such antibodies include monoclonal and polyclonal antibodies, single chain antibodies, chimeric antibodies, bifunctional/bispecific antibodies, humanized antibodies, human antibodies, and complementary determining region (CDR)-grafted antibodies, including compounds which include CDR and/or antigen-binding sequences, which specifically recognize a polypeptide of the invention. Preferred antibodies of the invention are human antibodies which are produced and identified according to methods described in WO93/11236, published June 20, 1993, which is incorporated herein by reference in its entirety. Antibody fragments, including Fab, Fab′, F(ab′)$_2$, and F$_v$, are also provided by the invention. The term "specific for" indicates that the variable regions of the antibodies of the invention recognize and bind CG122, CG179, CG95, CG121, CG162, CG27, CG153 or CG168 polypeptides exclusively (*i.e.*, able to distinguish a CG122 or CG179 polypeptide from other apolipoprotein polypeptides; CG95, CG121 or CG162 polypeptide from other lipase polypeptide; CG27, CG153 or CG168 polypeptide from other lipoprotein receptor polypeptide, despite sequence identity, homology, or similarity found in the family of polypeptides), but may also interact with other proteins (for example, *S. aureus* protein A or other antibodies in ELISA techniques) through interactions with sequences outside the variable region of the antibodies, and in particular, in the constant region of the molecule. Screening assays to determine binding specificity of an antibody of the invention are well known and routinely practiced in the art. For a comprehensive discussion of such assays, see Harlow et al. (Eds), Antibodies A Laboratory Manual; Cold Spring Harbor Laboratory; Cold Spring Harbor , NY (1988), Chapter 6. Antibodies that recognize and bind fragments of the CG122, CG179, CG95, CG121, CG162, CG27, CG153, or CG168 polypeptides of the invention are also contemplated, provided that the antibodies are first and foremost specific for, as defined above, CG122, CG179, CG95,

CG121, CG162, CG27, CG153 or CG168 polypeptides. As with antibodies that are specific for full length apolipoprotein polypeptides, antibodies of the invention that recognize CG122 or CG179 are those which can distinguish CG122 or CG179 polypeptides from the family of apolipoprotein polypeptides despite inherent sequence identity, homology, or similarity found in the family of proteins. As with antibodies that are specific for full length lipase polypeptides, antibodies of the invention that recognize CG95, CG121 or CG162 are those which can distinguish CG95, CG121 or CG162 polypeptides from the family of lipase polypeptides despite inherent sequence identity, homology, or similarity found in the family of proteins. As with antibodies that are specific for full length lipoprotein receptor polypeptides, antibodies of the invention that recognize CG27, CG153 or CG168 are those which can distinguish CG27, CG153 or CG168 polypeptides from the family of lipoprotein receptor polypeptides despite inherent sequence identity, homology, or similarity found in the family of proteins. Antibodies of the invention can be produced using any method well known and routinely practiced in the art.

Non-human antibodies may be humanized by any methods known in the art. In one method, the non-human CDRs are inserted into a human antibody or consensus antibody framework sequence. Further changes can then be introduced into the antibody framework to modulate affinity or immunogenicity.

Antibodies of the invention are useful for, for example, therapeutic purposes (by modulating activity of a polypeptide of the invention), diagnostic purposes to detect or quantitate a polypeptide of the invention, as well as purification of a polypeptide of the invention. Kits comprising an antibody of the invention for any of the purposes described herein are also comprehended. In general, a kit of the invention also includes a control antigen for which the antibody is immunospecific. The invention further provides a hybridoma that produces an antibody according to the invention. Antibodies of the invention are useful for detection and/or purification of the polypeptides of the invention.

Proteins of the invention may also be used to immunize animals to obtain polyclonal and monoclonal antibodies which specifically react with the protein. Such antibodies may be obtained using either the entire protein or fragments thereof as an immunogen. The peptide immunogens additionally may contain a cysteine residue at the carboxyl terminus, and are conjugated to a hapten such as keyhole limpet hemocyanin

(KLH). Methods for synthesizing such peptides are known in the art, for example, as in R. P. Merrifield, J. Amer. Chem. Soc. 85, 2149-2154 (1963); J. L. Krstenansky, et al., FEBS Lett. 211, 10 (1987). Monoclonal antibodies binding to the protein of the invention may be useful diagnostic agents for the immunodetection of the protein. Neutralizing monoclonal antibodies binding to the protein may also be useful therapeutics for both conditions associated with the protein and also in the treatment of some forms of cancer where abnormal expression of the protein is involved. In the case of cancerous cells or leukemic cells, neutralizing monoclonal antibodies against the protein may be useful in detecting and preventing the metastatic spread of the cancerous cells, which may be mediated by the protein. In general, techniques for preparing polyclonal and monoclonal antibodies as well as hybridomas capable of producing the desired antibody are well known in the art (Campbell, A.M., Monoclonal Antibodies Technology: Laboratory Techniques in Biochemistry and Molecular Biology, Elsevier Science Publishers, Amsterdam, The Netherlands (1984); St. Groth et al., J. Immunol. 35:1-21 (1990); Kohler and Milstein, Nature 256:495-497 (1975)), the trioma technique, the human B-cell hybridoma technique (Kozbor et al., Immunology Today 4:72 (1983); Cole et al., in Monoclonal Antibodies and Cancer Therapy, Alan R. Liss, Inc. (1985), pp. 77-96).

Any animal (mouse, rabbit, etc.) which is known to produce antibodies can be immunized with a peptide or polypeptide of the invention. Methods for immunization are well known in the art. Such methods include subcutaneous or intraperitoneal injection of the polypeptide. One skilled in the art will recognize that the amount of the protein encoded by the ORF of the present invention used for immunization will vary based on the animal which is immunized, the antigenicity of the peptide and the site of injection. The protein that is used as an immunogen may be modified or administered in an adjuvant in order to increase the protein's antigenicity. Methods of increasing the antigenicity of a protein are well known in the art and include, but are not limited to, coupling the antigen with a heterologous protein (such as globulin or β-galactosidase) or through the inclusion of an adjuvant during immunization.

For monoclonal antibodies, spleen cells from the immunized animals are removed, fused with myeloma cells, such as SP2/0-Ag14 myeloma cells, and allowed to become monoclonal antibody producing hybridoma cells. Any one of a number of methods well known in the art can be used to identify the hybridoma cell which produces an antibody

with the desired characteristics. These include screening the hybridomas with an ELISA assay, western blot analysis, or radioimmunoassay (Lutz et al., Exp. Cell Research. 175:109-124 (1988)). Hybridomas secreting the desired antibodies are cloned and the class and subclass is determined using procedures known in the art (Campbell, A.M., Monoclonal Antibody Technology: Laboratory Techniques in Biochemistry and Molecular Biology, Elsevier Science Publishers, Amsterdam, The Netherlands (1984)). Techniques described for the production of single chain antibodies (U.S. Patent 4,946,778) can be adapted to produce single chain antibodies to proteins of the present invention.

For polyclonal antibodies, antibody containing antiserum is isolated from the immunized animal and is screened for the presence of antibodies with the desired specificity using one of the above-described procedures. The present invention further provides the above-described antibodies in delectably labeled form. Antibodies can be delectably labeled through the use of radioisotopes, affinity labels (such as biotin, avidin, etc.), enzymatic labels (such as horseradish peroxidase, alkaline phosphatase, etc.) fluorescent labels (such as FITC or rhodamine, etc.), paramagnetic atoms, etc. Procedures for accomplishing such labeling are well-known in the art, for example, see (Sternberger, L.A. et al., J. Histochem. Cytochem. 18:315 (1970); Bayer, E.A. et al., Meth. Enzym. 62:308 (1979); Engval, E. et al., Immunol. 109:129 (1972); Goding, J.W. J. Immunol. Meth. 13:215 (1976)).

The labeled antibodies of the present invention can be used for *in vitro*, *in vivo*, and *in situ* assays to identify cells or tissues in which a fragment of the polypeptide of interest is expressed. The antibodies may also be used directly in therapies or other diagnostics. The present invention further provides the above-described antibodies immobilized on a solid support. Examples of such solid supports include plastics such as polycarbonate, complex carbohydrates such as agarose and Sepharose®, acrylic resins and such as polyacrylamide and latex beads. Techniques for coupling antibodies to such solid supports are well known in the art (Weir, D.M. et al., "Handbook of Experimental Immunology" 4th Ed., Blackwell Scientific Publications, Oxford, England, Chapter 10 (1986); Jacoby, W.D. et al., Meth. Enzym. 34 Academic Press, N.Y. (1974)). The immobilized antibodies of the present invention can be used for *in vitro*, *in vivo*, and *in*

*situ* assays as well as for immuno-affinity purification of the proteins of the present invention.

**9. COMPUTER READABLE SEQUENCES**

In one application of this embodiment, a nucleotide sequence of the present invention can be recorded on computer readable media. As used herein, "computer readable media" refers to any medium which can be read and accessed directly by a computer. Such media include, but are not limited to: magnetic storage media, such as floppy discs, hard disc storage medium, and magnetic tape; optical storage media such as CD-ROM; electrical storage media such as RAM and ROM; and hybrids of these categories such as magnetic/optical storage media. A skilled artisan can readily appreciate how any of the presently known computer readable mediums can be used to create a manufacture comprising computer readable medium having recorded thereon a nucleotide sequence of the present invention. As used herein, "recorded" refers to a process for storing information on computer readable medium. A skilled artisan can readily adopt any of the presently known methods for recording information on computer readable medium to generate manufactures comprising the nucleotide sequence information of the present invention.

A variety of data storage structures are available to a skilled artisan for creating a computer readable medium having recorded thereon a nucleotide sequence of the present invention. The choice of the data storage structure will generally be based on the means chosen to access the stored information. In addition, a variety of data processor programs and formats can be used to store the nucleotide sequence information of the present invention on computer readable medium. The sequence information can be represented in a word processing text file, formatted in commercially-available software such as WordPerfect and Microsoft Word, or represented in the form of an ASCII file, stored in a database application, such as DB2, Sybase, Oracle, or the like. A skilled artisan can readily adapt any number of data processor structuring formats (e.g. text file or database) in order to obtain computer readable medium having recorded thereon the nucleotide sequence information of the present invention. By providing the nucleotide sequence of SEQ ID NO: 1, 3, 5, 7, 9, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, 34, 36, 38, 40, 42 or 44; or a representative fragment thereof; or a nucleotide sequence at least 99.9%

identical to SEQ ID NO: 1, 3, 5, 7, 9, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, 34, 36, 38, 40, 42 or 44 in computer readable form, a skilled artisan can routinely access the sequence information for a variety of purposes. Computer software is publicly available which allows a skilled artisan to access sequence information provided in a computer readable medium. The examples which follow demonstrate how software which implements the BLAST (Altschul et al., J. Mol. Biol. 215:403-410 (1990)) and BLAZE (Brutlag et al., Comp. Chem. 17:203-207 (1993)) search algorithms on a Sybase system is used to identify open reading frames (ORFs) within a nucleic acid sequence. Such ORFs may be protein encoding fragments and may be useful in producing commercially important proteins such as enzymes used in fermentation reactions and in the production of commercially useful metabolites.

As used herein, "a computer-based system" refers to the hardware means, software means, and data storage means used to analyze the nucleotide sequence information of the present invention. The minimum hardware means of the computer-based systems of the present invention comprises a central processing unit (CPU), input means, output means, and data storage means. A skilled artisan can readily appreciate that any one of the currently available computer-based systems are suitable for use in the present invention. As stated above, the computer-based systems of the present invention comprise a data storage means having stored therein a nucleotide sequence of the present invention and the necessary hardware means and software means for supporting and implementing a search means. As used herein, "data storage means" refers to memory which can store nucleotide sequence information of the present invention, or a memory access means which can access manufactures having recorded thereon the nucleotide sequence information of the present invention.

As used herein, "search means" refers to one or more programs which are implemented on the computer-based system to compare a target sequence or target structural motif with the sequence information stored within the data storage means. Search means are used to identify fragments or regions of a known sequence which match a particular target sequence or target motif. A variety of known algorithms are disclosed publicly and a variety of commercially available software for conducting search means are and can be used in the computer-based systems of the present invention. Examples of such software includes, but is not limited to, MacPattern (EMBL), BLASTN and

BLASTA (NPOLYPEPTIDEIA). A skilled artisan can readily recognize that any one of the available algorithms or implementing software packages for conducting homology searches can be adapted for use in the present computer-based systems. As used herein, a "target sequence" can be any nucleic acid or amino acid sequence of six or more nucleotides or two or more amino acids. A skilled artisan can readily recognize that the longer a target sequence is, the less likely a target sequence will be present as a random occurrence in the database. The most preferred sequence length of a target sequence is from about 10 to 100 amino acids or from about 30 to 300 nucleotide residues. However, it is well recognized that searches for commercially important fragments, such as sequence fragments involved in gene expression and protein processing, may be of shorter length.

As used herein, "a target structural motif," or "target motif," refers to any rationally selected sequence or combination of sequences in which the sequence(s) are chosen based on a three-dimensional configuration which is formed upon the folding of the target motif. There are a variety of target motifs known in the art. Protein target motifs include, but are not limited to, enzyme active sites and signal sequences. Nucleic acid target motifs include, but are not limited to, promoter sequences, hairpin structures and inducible expression elements (protein binding sequences).



## 10. TRIPLE HELIX FORMATION

In addition, gene expression can be controlled through triple helix formation or antisense DNA or RNA, both of which methods are based on the binding of a polynucleotide sequence to DNA or RNA. Polynucleotides suitable for use in these methods are usually 20 to 40 bases in length and are designed to be complementary to a region of the gene involved in transcription (triple helix - see Lee et al., Nucl. Acids Res. 6:3073 (1979); Cooney et al., Science 15241:456 (1988); and Dervan et al., Science 251:1360 (1991)) or to the mRNA itself (antisense - Olmno, J. Neurochem. 56:560 (1991); Oligodeoxynucleotides as Antisense Inhibitors of Gene Expression, CRC Press, Boca Raton, FL (1988)). Triple helix- formation optimally results in a shut-off of RNA transcription from DNA, while antisense RNA hybridization blocks translation of an mRNA molecule into polypeptide. Both techniques have been demonstrated to be

effective in model systems. Information contained in the sequences of the present invention is necessary for the design of an antisense or triple helix oligonucleotide.

## 11. DIAGNOSTIC ASSAYS AND KITS

The present invention further provides methods to identify the presence or expression of one of the ORFs of the present invention, or homolog thereof, in a test sample, using a nucleic acid probe or antibodies of the present invention, optionally conjugated or otherwise associated with a suitable label.

In general, methods for detecting a polynucleotide of the invention can comprise contacting a sample with a compound that binds to and forms a complex with the polypeptide for a period sufficient to form the complex, and detecting the complex, so that if a complex is detected, a polypeptide of the invention is detected in the sample. In detail, such methods comprise incubating a test sample with one or more of the antibodies or one or more of nucleic acid probes of the present invention and assaying for binding of the nucleic acid probes or antibodies to components within the test sample.

Conditions for incubating a nucleic acid probe or antibody with a test sample vary. Incubation conditions depend on the format employed in the assay, the detection methods employed, and the type and nature of the nucleic acid probe or antibody used in the assay. One skilled in the art will recognize that any one of the commonly available hybridization, amplification or immunological assay formats can readily be adapted to employ the nucleic acid probes or antibodies of the present invention. Examples of such assays can be found in Chard, T., An Introduction to Radioimmunoassay and Related Techniques, Elsevier Science Publishers, Amsterdam, The Netherlands (1986); Bullock, G.R. et al., Techniques in Immunocytochemistry, Academic Press, Orlando, FL Vol. 1 (1982), Vol. 2 (1983), Vol. 3 (1985); Tijssen, P., Practice and Theory of immunoassays: Laboratory Techniques in Biochemistry and Molecular Biology, Elsevier Science Publishers, Amsterdam, The Netherlands (1985). The test samples of the present invention include cells, protein or membrane extracts of cells, or biological fluids such as sputum, blood, serum, plasma, or urine. The test sample used in the above-described method will vary based on the assay format, nature of the detection method and the tissues, cells or extracts used as the sample to be assayed. Methods for preparing protein

extracts or membrane extracts of cells are well known in the art and can be readily be adapted in order to obtain a sample which is compatible with the system utilized.

In another embodiment of the present invention, kits are provided which contain the necessary reagents to carry out the assays of the present invention. Specifically, the invention provides a compartment kit to receive, in close confinement, one or more containers which comprises: (a) a first container comprising one of the probes or antibodies of the present invention; and (b) one or more other containers comprising one or more of the following: wash reagents, reagents capable of detecting presence of a bound probe or antibody.

In detail, a compartment kit includes any kit in which reagents are contained in separate containers. Such containers include small glass containers, plastic containers or strips of plastic or paper. Such containers allows one to efficiently transfer reagents from one compartment to another compartment such that the samples and reagents are not cross-contaminated, and the agents or solutions of each container can be added in a quantitative fashion from one compartment to another. Such containers will include a container which will accept the test sample, a container which contains the antibodies used in the assay, containers which contain wash reagents (such as phosphate buffered saline, Tris-buffers, etc.), and containers which contain the reagents used to detect the bound antibody or probe. Types of detection reagents include labeled nucleic acid probes, labeled secondary antibodies, or in the alternative, if the primary antibody is labeled, the enzymatic, or antibody binding reagents which are capable of reacting with the labeled antibody. One skilled in the art will readily recognize that the disclosed probes and antibodies of the present invention can be readily incorporated into one of the established kit formats which are well known in the art.

**12. MEDICAL IMAGING**

The novel polypeptides of the invention are useful in medical imaging, e.g., imaging the site of infection, inflammation, and other sites expressing CG122 or CG179 apolipoprotein molecules; CG95, CG121 or CG162 lipase molecules; or CG27, CG153 or CG168 lipoprotein receptor molecules. See, e.g., Kunkel et al., U.S. Pat. NO. 5,413,778. Such methods involve chemical attachment of a labeling or imaging agent, administration

of the labeled polypeptide to a subject in a pharmaceutically acceptable carrier, and imaging the labeled polypeptide *in vivo* at the target site.

**13. SCREENING ASSAYS**

Using the isolated proteins and polynucleotides of the invention, the present invention further provides methods of obtaining and identifying agents which bind to a polypeptide encoded by the ORF from a polynucleotide of the invention to a specific domain of the polypeptide encoded by a polypeptide of the invention. In detail, said method comprises the steps of:

(a) contacting an agent with an isolated protein encoded by an ORF of the present invention, or nucleic acid of the invention; and

(b) determining whether the agent binds to said protein or said nucleic acid.

In general, therefore, such methods for identifying compounds that bind to a polynucleotide of the invention can comprise contacting a compound with a polynucleotide of the invention for a time sufficient to form a polynucleotide/compound complex, and detecting the complex, so that if a polynucleotide/compound complex is detected, a compound that binds to a polynucleotide of the invention is identified.

Likewise, in general, therefore, such methods for identifying compounds that bind to a polypeptide of the invention can comprise contacting a compound with a polypeptide of the invention for a time sufficient to form a polypeptide/compound complex, and detecting the complex, so that if a polypeptide/compound complex is detected, a compound that binds to a polynucleotide of the invention is identified.

Methods for identifying compounds that bind to a polypeptide of the invention can also comprise contacting a compound with a polypeptide of the invention in a cell for a time sufficient to form a polypeptide/compound complex, wherein the complex drives expression of a receptor gene sequence in the cell, and detecting the complex by detecting reporter gene sequence expression, so that if a polypeptide/compound complex is detected, a compound that binds a polypeptide of the invention is identified.

Compounds identified via such methods can include compounds which modulate the activity of a polypeptide of the invention (that is, increase or decrease its activity, relative to activity observed in the absence of the compound). Alternatively, compounds identified via such methods can include compounds which modulate the expression of a polynucleotide of the invention (that is, increase or decrease expression relative to expression levels observed in the absence of the compound). Compounds, such as compounds identified via the methods of the invention, can be tested using standard

assays well known to those of skill in the art for their ability to modulate activity/expression.

The agents screened in the above assay can be, but are not limited to, peptides, carbohydrates, vitamin derivatives, or other pharmaceutical agents. The agents can be selected and screened at random or rationally selected or designed using protein modeling techniques.

For random screening, agents such as peptides, carbohydrates, pharmaceutical agents and the like are selected at random and are assayed for their ability to bind to a protein encoded by an ORF of the present invention. Alternatively, agents may be rationally selected or designed. As used herein, an agent is said to be "rationally selected or designed" when the agent is chosen based on the configuration of the particular protein. For example, one skilled in the art can readily adapt currently available procedures to generate peptides, pharmaceutical agents and the like capable of binding to a specific peptide sequence in order to generate rationally designed antipeptide peptides, for example see Hurby et al., Application of Synthetic Peptides: Antisense Peptides," In Synthetic Peptides, A User's Guide, W.H Freeman, NY (1992), pp. 289-307, and Kaspczak et al., Biochemistry 28:9230-8 (1989), or pharmaceutical agents, or the like.

In addition to the foregoing, one class of agents of the present invention, as broadly described, can be used to control gene expression through binding to one of the ORFs or EMFs of the present invention. As described above, such agents can be randomly screened or rationally designed/selected. Targeting the ORF or EMF allows a skilled artisan to design sequence specific or element specific agents, modulating the expression of either a single ORF or multiple ORFs which rely on the same EMF for expression control. One class of DNA binding agents are agents which contain base residues which hybridize or form a triple helix formation by binding to DNA or RNA. Such agents can be based on the classic phosphodiester, ribonucleic acid backbone, or can be a variety of sulfhydryl or polymeric derivatives which have base attachment capacity.

Agents suitable for use in these methods usually contain 20 to 40 bases and are designed to be complementary to a region of the gene involved in transcription (triple helix - see Lee et al., Nucl. Acids Res. 6:3073 (1979); Cooney et al., Science 241:456 (1988); and Dervan et al., Science 251:1360 (1991)) or to the mRNA itself (antisense - Okano, J. Neurochem. 56:560 (1991); Oligodeoxynucleotides as Antisense Inhibitors of

Gene Expression, CRC Press, Boca Raton, FL (1988)). Triple helix- formation optimally results in a shut-off of RNA transcription from DNA, while antisense RNA hybridization blocks translation of an mRNA molecule into polypeptide. Both techniques have been demonstrated to be effective in model systems. Information contained in the sequences of the present invention is necessary for the design of an antisense or triple helix oligonucleotide and other DNA binding agents. Agents which bind to a protein encoded by one of the ORFs of the present invention can be used as a diagnostic agent, in the control of bacterial infection by modulating the activity of the protein encoded by the ORF. Agents which bind to a protein encoded by one of the ORFs of the present invention can be formulated using known techniques to generate a pharmaceutical composition.

## 14. USE OF NUCLEIC ACIDS AS PROBES

Another aspect of the subject invention is to provide for polypeptide-specific nucleic acid hybridization probes capable of hybridizing with naturally occurring nucleotide sequences. The hybridization probes of the subject invention may be derived from the nucleotide sequence of the SEQ ID NO: 1, 3, 5, 7, 9, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, 34, 36, 38, 40, 42 or 44. Because the corresponding gene is only expressed in a limited number of tissues, a hybridization probe derived from SEQ ID NO: 1, 3, 5, 7, 9, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, 34, 36, 38, 40, 42 or 44 can be used as an indicator of the presence of RNA of cell type of such a tissue in a sample.

Any suitable hybridization technique can be employed, such as, for example, *in situ* hybridization. PCR as described US Patent Nos 4,683,195 and 4,965,188 provides additional uses for oligonucleotides based upon the nucleotide sequences. Such probes used in PCR may be of recombinant origin, may be chemically synthesized, or a mixture of both. The probe will comprise a discrete nucleotide sequence for the detection of identical sequences or a degenerate pool of possible sequences for identification of closely related genomic sequences.

Other means for producing specific hybridization probes for nucleic acids include the cloning of nucleic acid sequences into vectors for the production of mRNA probes. Such vectors are known in the art and are commercially available and may be used to synthesize RNA probes *in vitro* by means of the addition of the appropriate RNA

polymerase as T7 or SP6 RNA polymerase and the appropriate radioactively labeled nucleotides. The nucleotide sequences may be used to construct hybridization probes for mapping their respective genomic sequences. The nucleotide sequence provided herein may be mapped to a chromosome or specific regions of a chromosome using well known genetic and/or chromosomal mapping techniques. These techniques include *in situ* hybridization, linkage analysis against known chromosomal markers, hybridization screening with libraries or flow-sorted chromosomal preparations specific to known chromosomes, and the like. The technique of fluorescent *in situ* hybridization of chromosome spreads has been described, among other places, in Verma et al (1988) Human Chromosomes: A Manual of Basic Techniques, Pergamon Press, New York NY.

Fluorescent *in situ* hybridization of chromosomal preparations and other physical chromosome mapping techniques may be correlated with additional genetic map data. Examples of genetic map data can be found in the 1994 Genome Issue of Science (265:1981f). Correlation between the location of a nucleic acid on a physical chromosomal map and a specific disease (or predisposition to a specific disease) may help delimit the region of DNA associated with that genetic disease. The nucleotide sequences of the subject invention may be used to detect differences in gene sequences between normal, carrier or affected individuals. The nucleotide sequence may be used to produce purified polypeptides using well known methods of recombinant DNA technology. Among the many publications that teach methods for the expression of genes after they have been isolated is Goeddel (1990) Gene Expression Technology, Methods and Enzymology, Vol 185, Academic Press, San Diego. Polypeptides may be expressed in a variety of host cells, either prokaryotic or eukaryotic. Host cells may be from the same species from which a particular polypeptide nucleotide sequence was isolated or from a different species. Advantages of producing polypeptides by recombinant DNA technology include obtaining adequate amounts of the protein for purification and the availability of simplified purification procedures.

Each sequence so obtained was compared to sequences in GenBank using a search algorithm developed by Applied Biosystems and incorporated into the INHERIT™ 670 Sequence Analysis System. In this algorithm, Pattern Specification Language (developed by TRW Inc., Los Angeles, CA) was used to determine regions of homology. The three parameters that determine how the sequence comparisons run were window size, window

offset, and error tolerance. Using a combination of these three parameters, the DNA database was searched for sequences containing regions of homology to the query sequence, and the appropriate sequences were scored with an initial value. Subsequently, these homologous regions were examined using dot matrix homology plots to distinguish regions of homology from chance matches. Smith-Waterman alignments were used to display the results of the homology search. Peptide and protein sequence homologies were ascertained using the INHERIT™ 670 Sequence Analysis System in a way similar to that used in DNA sequence homologies. Pattern Specification Language and parameter windows were used to search protein databases for sequences containing regions of homology that were scored with an initial value. Dot-matrix homology plots were examined to distinguish regions of significant homology from chance matches.

Alternatively, BLAST, which stands for Basic Local Alignment Search Tool, is used to search for local sequence alignments (Altschul SF (1993) J Mol Evol 36:290-300; Altschul, SF et al (1990) J Mol Biol 215:403-10). BLAST produces alignments of both nucleotide and amino acid sequences to determine sequence similarity. Because of the local nature of the alignments, BLAST is especially useful in determining exact matches or in identifying homologs. Whereas it is ideal for matches which do not contain gaps, it is inappropriate for performing motif-style searching. The fundamental unit of BLAST algorithm output is the High-scoring Segment Pair (HSP). An HSP consists of two sequence fragments of arbitrary but equal lengths whose alignment is locally maximal and for which the alignment score meets or exceeds a threshold or cutoff score set by the user. The BLAST approach is to look for HSPs between a query sequence and a database sequence, to evaluate the statistical significance of any matches found, and to report only those matches which satisfy the user-selected threshold of significance. The parameter E establishes the statistically significant threshold for reporting database sequence matches. E is interpreted as the upper bound of the expected frequency of chance occurrence of an HSP (or set of HSPs) within the context of the entire database search. Any database sequence whose match satisfies E is reported in the program output.

**14.1 Preparation of Sequencing Chips and Arrays**

A basic example is using 6-mers attached to 50 micron surfaces to give a chip with dimensions of 3 x 3 mm which can be combined to give an array of 20 x 20 cm. Another example is using 9-mer oligonucleotides attached to 10 x 10 microns surface to

create a 9-mer chip, with dimensions of 5 x 5 mm. 4000 units of such chips may be used to create a 30 x 30 cm array. In an array in which 4,000 to 16,000 oligochips are arranged into a square array. A plate, or collection of tubes, as also depicted, may be packaged with the array as part of the sequencing kit.

The arrays may be separated physically from each other or by hydrophobic surfaces. One possible way to utilize the hydrophobic strip separation is to use technology such as the Iso-Grid Microbiology System produced by QA Laboratories, Toronto, Canada.

Hydrophobic grid membrane filters (HGMF) have been in use in analytical food microbiology for about a decade where they exhibit unique attractions of extended numerical range and automated counting of colonies. One commercially-available grid is ISO-GRID™ from QA Laboratories Ltd. (Toronto, Cánada) which consists of a square (60 x 60 cm) of polysulfone polymer (Gelman Tuffryn HT-450, 0.45u pore size) on which is printed a black hydrophobic ink grid consisting of 1600 (40 x 40) square cells. HGMF have previously been inoculated with bacterial suspensions by vacuum filtration and incubated on the differential or selective media of choice.

Because the microbial growth is confined to grid cells of known position and size on the membrane, the HGMF functions more like an MPN apparatus than a conventional plate or membrane filter. Peterkin *et al.* (1987) reported that these HGMFs can be used to propagate and store genomic libraries when used with a HGMF replicator. One such instrument replicates growth from each of the 1600 cells of the ISO-GRID and enables many copies of the master HGMF to be made (Peterkin *et al.,* 1987).

Sharpe *et al.* (1989) also used ISO-GRID HGMF form QA Laboratories and an automated HGMF counter (MI-100 Interpreter) and RP-100 Replicator. They reported a technique for maintaining and screening many microbial cultures.

Peterkin and colleagues later described a method for screening DNA probes using the hydrophobic grid-membrane filter (Peterkin *et al.*, 1989). These authors reported methods for effective colony hybridization directly on HGMFs. Previously, poor results had been obtained due to the low DNA binding capacity of the epoxysulfone polymer on which the HGMFs are printed. However, Peterkin *et al.* (1989) reported that the binding of DNA to the surface of the membrane was improved by treating the replicated and incubated HGMF with polyethyleneimine, a polycation, prior to contact with DNA. Although this

early work uses cellular DNA attachment, and has a different objective to the present invention, the methodology described may be readily adapted for Format 3 SBH.

In order to identify useful sequences rapidly, Peterkin *et al.* (1989) used radiolabeled plasmid DNA from various clones and tested its specificity against the DNA on the prepared HGMFs. In this way, DNA from recombinant plasmids was rapidly screened by colony hybridization against 100 organisms on HGMF replicates which can be easily and reproducibly prepared.

Manipulation with small (2-3 mm) chips, and parallel execution of thousands of the reactions. The solution of the invention is to keep the chips and the probes in the corresponding arrays. In one example, chips containing 250,000 9-mers are synthesized on a silicon wafer in the form of 8 x 8 mM plates (15 uM/oligonucleotide, Pease et al., 1994) arrayed in 8 x 12 format (96 chips) with a 1 mM groove in between. Probes are added either by multichannel pipette or pin array, one probe on one chip. To score all 4000 6-mers, 42 chip arrays have to be used, either using different ones, or by reusing one set of chip arrays several times.

In the above case, using the earlier nomenclature of the application, F=9; P=6; and F +P = 15. Chips may have probes of formula BxNn, where x is a number of specified bases B; and n is a number of non-specified bases, so that x = 4 to 10 and n = 1 to 4. To achieve more efficient hybridization, and to avoid potential influence of any support oligonucleotides, the specified bases can be surrounded by unspecified bases, thus represented by a formula such as (N)nBx(N)m.

### 14.2 Preparation of Support Bound Oligonucleotides

Oligonucleotides, i.e., small nucleic acid segments, may be readily prepared by, for example, directly synthesizing the oligonucleotide by chemical means, as is commonly practiced using an automated oligonucleotide synthesizer.

Support bound oligonucleotides may be prepared by any of the methods known to those of skill in the art using any suitable support such as glass, polystyrene or Teflon. One strategy is to precisely spot oligonucleotides synthesized by standard synthesizers. Immobilization can be achieved using passive adsorption (Inouye & Hondo, 1990); using UV light (Nagata *et al.,* 1985; Dahlen *et al.,* 1987; Morriey & Collins, 1989) or by covalent

binding of base modified DNA (Keller *et al.,* 1988; 1989); all references being specifically incorporated herein.

Another strategy that may be employed is the use of the strong biotin-streptavidin interaction as a linker. For example, Broude *et al.* (1994) describe the use of Biotinylated probes, although these are duplex probes, that are immobilized on streptavidin-coated magnetic beads. Streptavidin-coated beads may be purchased from Dynal, Oslo. Of course, this same linking chemistry is applicable to coating any surface with streptavidin. Biotinylated probes may be purchased from various sources, such as, e.g., Operon Technologies (Alameda, CA).

Nunc Laboratories (Naperville, IL) is also selling suitable material that could be used. Nunc Laboratories have developed a method by which DNA can be covalently bound to the microwell surface termed Covalink NH. CovaLink NH is a polystyrene surface grafted with secondary amino groups (>NH) that serve as bridge-heads for further covalent coupling. CovaLink Modules may be purchased from Nunc Laboratories. DNA molecules may be bound to CovaLink exclusively at the 5'-end by a phosphoramidate bond, allowing immobilization of more than 1 pmol of DNA (Rasmussen *et al.*, 1991).

The use of CovaLink NH strips for covalent binding of DNA molecules at the 5'-end has been described (Rasmussen et al., 1991). In this technology, a phosphoramidate bond is employed (Chu et al., 1983). This is beneficial as immobilization using only a single covalent bond is preferred. The phosphoramidate bond joins the DNA to the CovaLink NH secondary amino groups that are positioned at the end of spacer arms covalently grafted onto the polystyrene surface through a 2 nm long spacer arm. To link an oligonucleotide to CovaLink NH via an phosphoramidate bond, the oligonucleotide terminus must have a 5'-end phosphate group. It is, perhaps, even possible for biotin to be covalently bound to CovaLink and then streptavidin used to bind the probes.

More specifically, the linkage method includes dissolving DNA in water (7.5 ng/ul) and denaturing for 10 min. at 95°C and cooling on ice for 10 min. Ice-cold 0.1 M 1-methylimidazole, pH 7.0 (1-MeIm$_7$), is then added to a final concentration of 10 mM 1-MeIm$_7$. A ss DNA solution is then dispensed into CovaLink NH strips (75 ul/well) standing on ice.

Carbodiimide 0.2 M 1-ethyl-3-(3-dimethylaminopropyl)-carbodiimide (EDC), dissolved in 10 mM 1-MeIm$_7$, is made fresh and 25 ul added per well. The strips are

incubated for 5 hours at 50°C. After incubation the strips are washed using, e.g., Nunc-Immuno Wash; first the wells are washed 3 times, then they are soaked with washing solution for 5 min., and finally they are washed 3 times (where in the washing solution is 0.4 N NaOH, 0.25% SDS heated to 50°C).

It is contemplated that a further suitable method for use with the present invention is that described in PCT Patent Application WO 90/03382 (Southern & Maskos), incorporated herein by reference. This method of preparing an oligonucleotide bound to a support involves attaching a nucleoside 3'-reagent through the phosphate group by a covalent phosphodiester link to aliphatic hydroxyl groups carried by the support. The oligonucleotide is then synthesized on the supported nucleoside and protecting groups removed from the synthetic oligonucleotide chain under standard conditions that do not cleave the oligonucleotide from the support. Suitable reagents include nucleoside phosphoramidite and nucleoside hydrogen phosphorate.

An on-chip strategy for the preparation of DNA probe for the preparation of DNA probe arrays may be employed. For example, addressable laser-activated photodeprotection may be employed in the chemical synthesis of oligonucleotides directly on a glass surface, as described by Fodor *et al.* (1991), incorporated herein by reference. Probes may also be immobilized on nylon supports as described by Van Ness *et al.* (1991); or linked to Teflon using the method of Duncan & Cavalier (1988); all references being specifically incorporated herein.

To link an oligonucleotide to a nylon support, as described by Van Ness *et al.* (1991), requires activation of the nylon surface via alkylation and selective activation of the 5'-amine of oligonucleotides with cyanuric chloride.

One particular way to prepare support bound oligonucleotides is to utilize the light-generated synthesis described by Pease *et al.*, (1994, incorporated herein by reference). These authors used current photolithographic techniques to generate arrays of immobilized oligonucleotide probes (DNA chips). These methods, in which light is used to direct the synthesis of oligonucleotide probes in high-density, miniaturized arrays, utilize photolabile 5'-protected *N*-acyl-deoxynucleoside phosphoramidites, surface linker chemistry and versatile combinatorial synthesis strategies. A matrix of 256 spatially defined oligonucleotide probes may be generated in this manner and then used in the advantageous Format 3 sequencing, as described herein.

**14.3 Preparation of Nucleic Acid Fragments**

The nucleic acids to be sequenced may be obtained from any appropriate source, such as cDNAs, genomic DNA, chromosomal DNA, microdissected chromosome bands, cosmid or YAC inserts, and RNA, including mRNA without any amplification steps. For example, Sambrook *et al.* (1989) describes three protocols for the isolation of high molecular weight DNA from mammalian cells (p. 9.14-9.23).

DNA fragments may be prepared as clones in M13, plasmid or lambda vectors and/or prepared directly from genomic DNA or cDNA by PCR or other amplification methods. Samples may be prepared or dispensed in multiwell plates. About 100-1000 ng of DNA samples may be prepared in 2-500 ml of final volume.

The nucleic acids would then be fragmented by any of the methods known to those of skill in the art including, for example, using restriction enzymes as described at 9.24-9.28 of Sambrook *et al.* (1989), shearing by ultrasound and NaOH treatment.

Low pressure shearing is also appropriate, as described by Schriefer *et al.* (1990, incorporated herein by reference). In this method, DNA samples are passed through a small French pressure cell at a variety of low to intermediate pressures. A lever device allows controlled application of low to intermediate pressures to the cell. The results of these studies indicate that low-pressure shearing is a useful alternative to sonic and enzymatic DNA fragmentation methods.

One particularly suitable way for fragmenting DNA is contemplated to be that using the two base recognition endonuclease, *Cvi*JI, described by Fitzgerald *et al.* (1992). These authors described an approach for the rapid fragmentation and fractionation of DNA into particular sizes that they contemplated to be suitable for shotgun cloning and sequencing. The present inventor envisions that this will also be particularly useful for generating random, but relatively small, fragments of DNA for use in the present sequencing technology.

The restriction endonuclease *Cvi*JI normally cleaves the recognition sequence PuGCPy between the G and C to leave blunt ends. Atypical reaction conditions, which alter the specificity of this enzyme (*Cvi*JI**), yield a quasi-random distribution of DNA fragments form the small molecule pUC19 (2688 base pairs). Fitzgerald *et al.* (1992) quantitatively evaluated the randomness of this fragmentation strategy, using a *Cvi*JI** digest of pUC19 that was size fractionated by a rapid gel filtration method and directly

ligated, without end repair, to a lac Z minus M13 cloning vector. Sequence analysis of 76 clones showed that *Cvi*JI** restricts pyGCPy and PuGCPu, in addition to PuGCPy sites, and that new sequence data is accumulated at a rate consistent with random fragmentation.

As reported in the literature, advantages of this approach compared to sonication and agarose gel fractionation include: smaller amounts of DNA are required (0.2-0.5 ug instead of 2-5 ug); and fewer steps are involved (no preligation, end repair, chemical extraction, or agarose gel electrophoresis and elution are needed). These advantages are also proposed to be of use when preparing DNA for sequencing by Format 3.

Irrespective of the manner in which the nucleic acid fragments are obtained or prepared, it is important to denature the DNA to give single stranded pieces available for hybridization. This is achieved by incubating the DNA solution for 2-5 minutes at 80-90°C. The solution is then cooled quickly to 2°C to prevent renaturation of the DNA fragments before they are contacted with the chip. Phosphate groups must also be removed from genomic DNA by methods known in the art.

### 14.4 Preparation of DNA Arrays

Arrays may be prepared by spotting DNA samples on a support such as a nylon membrane. Spotting may be performed by using arrays of metal pins (the positions of which correspond to an array of wells in a microtiter plate) to repeated by transfer of about 20 nl of a DNA solution to a nylon membrane. By offset printing, a density of dots higher than the density of the wells is achieved. One to 25 dots may be accommodated in 1 $mm^2$, depending on the type of label used. By avoiding spotting in some preselected number of rows and columns, separate subsets (subarrays) may be formed. Samples in one subarray may be the same genomic segment of DNA (or the same gene) from different individuals, or may be different, overlapped genomic clones. Each of the subarrays may represent replica spotting of the same samples. In one example, a selected gene segment may be amplified from 64 patients. For each patient, the amplified gene segment may be in one 96-well plate (all 96 wells containing the same sample). A plate for each of the 64 patients is prepared. By using a 96-pin device, all samples may be spotted on one 8 x 12 cm membrane. Subarrays may contain 64 samples, one from each patient. Where the 96 subarrays are identical, the dot span may be 1 $mm^2$ and there may be a 1 mm space between subarrays.

Another approach is to use membranes or plates (available from NUNC, Naperville, Illinois) which may be partitioned by physical spacers e.g. a plastic grid molded over the membrane, the grid being similar to the sort of membrane applied to the bottom of multiwell plates, or hydrophobic strips. A fixed physical spacer is not preferred for imaging by exposure to flat phosphor-storage screens or x-ray films.

**14.5 Sequence Comparisons**

Preferred identity and/or similarity are designed to give the largest match between the sequences tested. Methods to determine identity and similarity are codified in publicly available computer programs including, but are not limited to, the GCG program package, including GAP (Devereux, J., et al., Nucleic Acids Research 12(1):387 (1984); Genetics Computer Group, University of Wisconsin, Madison, WI), BLASTP, BLASTN, BLASTX, and FASTA (Atschul, S.F. et al., J. Molec. Biol. 215:403-410 (1990). The BLAST X program is publicly available from the National Center for Biotechnology Information (NCBI) and other sources (BLAST Manual, Altschul, S., et al. NCB NLM NIH Bethesda, MD 20894; Altschul, S., et al., J. Mol. Biol. 215:403-410 (1990). The preferred computer program is FASTA version 3, specifically the FASTy program within the FASTA program package. Another preferred algorithm is the well known Smith Waterman algorithm which can also be used to determine identity.

Sequences can be compared to sequences in GenBank using a search algorithm developed by Applied Biosystems and incorporated into the INHERIT™ 670 Sequence Analysis System. In this algorithm, Pattern Specification Language (developed by TRW Inc., Los Angeles, CA) is used to determine regions of homology. The three parameters that determine how the sequence comparisons run are window size, window offset, and error tolerance. Using a combination of these three parameters, the DNA database can be searched for sequences containing regions of homology to the query sequence, and the appropriate sequences scored with an initial value. Subsequently, these homologous regions are examined using dot matrix homology plots to distinguish regions of homology from chance matches. Smith-Waterman alignments can be used to display the results of the homology search. Peptide and protein sequence homologies can be ascertained using the INHERIT™ 670 Sequence Analysis System in a way similar to that used in DNA sequence homologies. Pattern Specification Language and parameter windows are used to search

protein databases for sequences containing regions of homology that were scored with an initial value. Dot-matrix homology plots can be examined to distinguish regions of significant homology from chance matches.

Alternatively, BLAST, which stands for Basic Local Alignment Search Tool, is used to search for local sequence alignments (Altschul SF (1993) J Mol Evol 36:290-300; Altschul, SF et al (1990) J Mol Biol 215:403-10). BLAST produces alignments of both nucleotide and amino acid sequences to determine sequence similarity. Because of the local nature of the alignments, BLAST is especially useful in determining exact matches or in identifying homologs. Whereas it is ideal for matches which do not contain gaps, it is inappropriate for performing motif-style searching. The fundamental unit of BLAST algorithm output is the High-scoring Segment Pair (HSP). An HSP consists of two sequence fragments of arbitrary but equal lengths whose alignment is locally maximal and for which the alignment score meets or exceeds a threshold or cutoff score set by the user. The BLAST approach is to look for HSPs between a query sequence and a database sequence, to evaluate the statistical significance of any matches found, and to report only those matches which satisfy the user-selected threshold of significance. The parameter E establishes the statistically significant threshold for reporting database sequence matches. E is interpreted as the upper bound of the expected frequency of chance occurrence of an HSP (or set of HSPs) within the context of the entire database search.

The present invention is illustrated in the following examples. Upon consideration of the present disclosure, one of skill in the art will appreciate that many other embodiments and variations may be made in the scope of the present invention. Accordingly, it is intended that the broader aspects of the present invention not be limited to the disclosure of the following examples.

## EXAMPLE 1

### Cloning of Apolipoprotein, Lipase, and Lipoprotein Receptor cDNAs

Novel nucleic acids were obtained from various cDNA libraries (prepared from human mRNA purchased from Invitrogen, San Diego, CA) using standard PCR, sequencing by hybridization (SBH) sequence signature analysis and Sanger sequencing techniques. The inserts of the library were amplified with PCR using primers specific for pSport1 (GIBCO BRL, Grand Island, NY) vector sequences which flank the inserts. These samples were

spotted onto nylon membranes and hybridized with oligonucleotide probes to give sequence signatures. The clones were clustered into groups of similar or identical sequences, and single representative clones were selected from each group for gel sequencing. The 5' sequence of the amplified inserts was then deduced using the reverse M13 sequencing primer in a typical Sanger sequencing protocol. PCR products were purified and subjected to flourescent dye terminator cycle sequencing. Single pass gel sequencing was done using a 377 Applied Biosystems (ABI) sequencer.

Sequence analysis identified a polynucleotides encoding novel polypeptides designated CG122, CG179, CG95, CG121, CG162, CG27, CG153, and CG168. The 5' sequence was determined as described in Example 2.

**EXAMPLE 2**

**5' RACE Extension of Genes**

5' RACE reactions were performed using pairs of nested gene-specfic primers (GSP) and vector primers (VP) in sequential PCR reactions on a panel of cDNA libraries. The cDNA libraries used for RACE were prepared from mRNA using a random-primed, 5' capture method to enrich for the 5' ends of genes (Carninci et al, Genomics, 37, 327-336, 1996) and cloned into the pSPORT vector (BRL Life Technologies) previously digested with NotI and SalI. The human mRNAs (Invitrogen) included message from adult brain, adult thymus, fetal muscle, fetal skin, fetal heart, fetal brain, fetal spleen, fetal liver, and fetal lung. In addition, adaptor-ligated cDNA pools (Marathon cDNAs, Clontech) made from human fetal kidney, fetal brain and adult ovary mRNAs were used in the RACE experiments.

Generally, in the first reaction, a first GSP ($T_m$ ~80°C) and VP ($T_m$ ~72°C) are mixed in a 5:1 ratio. Touchdown PCR was carried out as follows: an initial incubation at 96°C for one minute, followed by five cycles of 96°C for 30 seconds and 72°C for four minutes; five cycles of 96°C for 30 seconds and 70°C for four minutes; and 15 cycles of 96°C for 30 seconds and 68°C for four minutes. The products of the first reaction were diluted 1:20 and used as template for the second reaction. Second nested GSP and VP (both $T_m$ ~60°C) were mixed in a 1:1 ratio and PCR was carried out as follows: an initial incubation at 96°C for one minute; and 30 cycles of 96°C for 30 seconds, 55°C for 30 seconds, and 72°C for 90 seconds. This step was sometimes repeated with a third or more nested GSP and VP primer. Final RACE products were separated and identified using

agarose gel electrophoresis. Selected fragments were subcloned into a TA cloning vector and the inserts were sequenced.

## EXAMPLE 3

### Tissue Expression Study

PCR Analysis

Gene expression of the polypeptides of the invention is analyzed using a semi-quantitative PCR-based technique. A panel of cDNA libraries derived from human tissue (from Clontech and Invitrogen) is screened with gene specific primers to examine the mRNA expression of the gene in human tissues and cell types. PCR assays (For example, 94 °C for 30 sec., 58 °C for 30 sec., 72 °C for 30 sec., for 30 cycles) are performed with 20 ng of cDNA derived from human tissues and cell lines and 10 picomoles of the appropriate gene-specific primers. The PCR product is identified through gel electrophoresis. Amplified products are separated on an agarose gel, transferred and chemically linked to a nylon filter. The filter is then hybridized with a radioactively labeled ($^{33}P\alpha$-dCTP) double-stranded probe generated from the full-length sequence using a Klenow polymerase, random prime method. The filters are washed (high stringency) and used to expose a phosphorimaging screen for several hours. Bands of the appropriate size indicate the presence of cDNA sequences in a specific library, and thus mRNA expression in the corresponding cell type or tissue.

Expression analysis can also be conducted using Northern blot techniques.

## EXAMPLE 4

### Chromosomal Localization Study

Chromosome mapping technologies allow investigators to link genes to specific regions of chromosomes. Chromosomal mapping is performed using the NIGMS human/rodent somatic cell hybrid mapping panel as described by Drwinga, H. L. et al., Genomics, 16, 311- 314, 1993 (human/rodent somatic cell hybrid mapping panel #2 purchased from the Coriell Institute for Medical Research, Camden, New Jersey). 60 ng of DNA from each sample in the panel is used as template, and 10 picomoles of the appropriate gene-specific oligonucleotides are used as primers in a PCR assay (for example, 94°C for 30 sec., 58°C for 30 sec., 72°C for 30 sec., for 30 cycles). PCR products were

analyzed by gel electrophoresis. The genomic PCR product is detected in a human/rodent somatic cell hybrid DNA containing a specific human chromosome.

## EXAMPLE 5

### Expression of Polypeptides in E. coli

A nucleic acid sequence of the invention is expressed in E. coli by subcloning the entire coding region into a prokaryotic expression vector. The expression vector (pQE16) used is from the QIAexpression® prokaryotic protein expression system (QIAGEN). The features of this vector that make it useful for protein expression include: an efficient promoter (phage T5) to drive transcription; expression control provided by the lac operator system, which can be induced by addition of IPTG (isopropyl-β-D-thiogalactopyranoside), and an encoded $His_6$ tag. The latter is a stretch of 6 histidine amino acid residues which can bind very tightly to a nickel atom. The vector can be used to express a recombinant protein with a $His_6$ tag fused to its carboxyl terminus, allowing rapid and efficient purification using Ni-coupled affinity columns.

PCR is used to amplify the coding region which is then ligated into digested pQE16 vector. The ligation product is transformed by electroporation into electrocompetent E.coli cells (strain M15[pREP4] from QIAGEN), and the transformed cells are plated on ampicillin-containing plates. Colonies are screened for the correct insert in the proper orientation using a PCR reaction employing a gene-specific primer and a vector-specific primer. Positives are then sequenced to ensure correct orientation and sequence. To express cytokine receptor polypeptides, a colony containing a correct recombinant clone is inoculated into L-Broth containing 100 μg/ml of ampicillin, 25 μg/ml of kanamycin, and the culture was allowed to grow overnight at 37°C. The saturated culture is then diluted 20-fold in the same medium and allowed to grow to an optical density at 600 nm of 0.5. At this point, IPTG is added to a final concentration of 1 mM to induce protein expression. The culture is allowed to grow for 5 more hours, and then the cells are harvested by centrifugation at 3000 x g for 15 minutes.

The resultant pellet is lysed using a mild, nonionic detergent in 20 mM Tris HCl (pH 7.5) (B-PER™ Reagent from Pierce), or by sonication until the turbid cell suspension turned translucent. The lysate obtained is further purified using a nickel containing column (Ni-NTA spin column from QIAGEN) under non-denaturing conditions. Briefly, the lysate

is brought up to 300 mM NaCl and 10 mM imidazole and centrifuged at 700 x g through the spin column to allow the His-tagged recombinant protein to bind to the nickel column. The column is then washed twice with Wash Buffer (50 mM $NaH_2PO_4$, pH 8.0; 300 mM NaCl; 20 mM imidazole) and is eluted with Elution Buffer (50 mM $NaH_2PO_4$, pH 8.0; 300 mM NaCl; 250 mM imidazole). All the above procedures are performed at 4°C. The presence of a purified protein of the predicted size is confirmed with SDS-PAGE.

## EXAMPLE 6

### Evaluation of Activities *In Vitro and In Vivo*

The activity of the polypeptides of the invention is assayed by monitoring the effect of such polypeptides on the activity of various signal transduction pathways. One commercially available system for monitoring signal transduction is the Dual-Luciferase™ Reporter Assay System (Promega Corp., Madison, WI). Briefly, mammalian cells are co-transfected with (1) a construct expressing the lipoprotein receptor polypeptide to be tested (*e.g.* CG27, CG153, CG168; or an active fragment; or an active fusion protein), (2) a first reporter construct utilizing a constitutive promoter (as a control for monitoring transfection efficiency), and (3) a second reporter construct that is dependent on a transcription factor or an enhancer element involved in the signal transduction pathway of interest (which serves to monitor the activity of one of several signal transduction pathways).

Various second reporter constructs are available in both cis- and trans-configurations (from, e.g., Stratagene, La Jolla, CA). The trans-configuration involves two constructs, and is used to monitor direct or indirect effects on signal transduction pathways which activate one of several transcription factors. Second reporter constructs for the following transcription factors are currently available from Stratagene: the Elk1 transcription factor for the mitogen-activated protein kinase (MAPK) signaling pathway, the c-Jun transcription factor for the c-Jun N-terminal kinase (JNK) signaling pathway, the CREB transcription factor for the cAMP-dependent kinase (PKA) signaling pathway, the CHOP transcription factor for the p38 kinase signaling pathway, and the c-Fos and ATF2 transcription factors. The cis-configuration is used to monitor direct or indirect effects on six different enhancer elements. Second reporter constructs for the following enhancer elements are currently available from Stratagene: AP-1, CRE, NF-kappaB, SRE, SRF and

p53. Other similar set of constructs may be prepared to monitor other transcription factors and enhancer elements known in the art.

Lipoproteins, or other exogenous ligand, either alone or in combination with other lipoproteins can be added to the transfected cells to determine the effects on candidate signal transduction pathways. Comparison of the effects on different pathways will show specificity of the lipoprotein receptor's biological effects.

In addition, this system can be used to screen libraries for small molecule drug candidates or lead compounds that disrupt or enhance the effects of the lipoprotein receptor.

## EXAMPLE 7

## Extension of sequences and identification of variants

Some of the novel nucleic acids of the present invention were assembled from sequences that were obtained from a cDNA library by methods described in Example 1 above, and in some cases sequences obtained from one or more public databases. The nucleic acids of SEQ ID NO: 16-42 were assembled using an EST sequence as a seed. Then a recursive algorithm was used to extend some of the seed ESTs into an extended assemblage, by pulling additional sequences from different databases (i.e., Hyseq's database containing EST sequences, dbEST version 122, gb pri 122, and UniGene version 122, Genseq 200105 (Derwent), and Genscan, Genemark and Hyseq gene predictions on human genomic sequence from the human genome project) that belong to this assemblage. The algorithm terminated when there was no additional sequences from the above databases that would extend the assemblage. Inclusion of component sequences into the assemblage was based on a BLASTN hit to the extending assemblage with BLAST score greater than 300 and percent identity greater than 95%.

Using PHRAP (Univ. of Washington) or CAP4 (Paracel), full-length gene cDNA sequences and their corresponding protein sequences were generated from the assemblage. Any frame shifts and incorrect stop codons were corrected by hand editing. During editing, the sequence was checked using FASTXY algorithm against Genbank (i.e., dbEST version 122, gb pri 122, UniGene version 122, Genpept release 122). Other computer programs which may have been used in the editing process were phredPhrap and Consed ((University of Washington) and ed-ready, ed-ext and gc-zip-2 (Hyseq, Inc.)).

## EXAMPLE 8

### *In vitro* and *in vivo* activity

A protein of the invention may also be tested for activity *in vitro* or *in vivo* using any assays known in the art. For example, assays for HDL, LDL or VLDL uptake or catabolism, beta-amyloid precursor protein (APP) uptake or catabolism, assays for anti-viral effects e.g. on virus assembly or budding, assays for effect on smooth muscle cell cultures, and animal models of atherosclerotic lesions induced by a variety of insults, e.g. high cholesterol diet or endothelial denudation, are described in Perrey et al., *Atherosclerosis*, 154:51-60 (2001), Kanaki et al., *Arteriosclerosis, Thrombosis and Vascular Biol.*, 19:2687 (1999), Kounnas et al., *Cell*, 82:331-340 (1995), and Fischer et al., *Science*, 262:250 (1993), the disclosures of all of which are incorporated by reference in their entirety.

The present invention is not to be limited in scope by the exemplified embodiments which are intended as illustrations of single aspects of the invention, and compositions and methods which are functionally equivalent are within the scope of the invention. Indeed, numerous modifications and variations in the practice of the invention are expected to occur to those skilled in the art upon consideration of the present preferred embodiments. Consequently, the only limitations which should be placed upon the scope of the invention are those which appear in the appended claims. All references cited within the body of the instant specification are hereby incorporated by reference in their entirety.